Methods in Molecular Biology 2024

Springer Protocols



Kelly M. Fulton
Susan M. Twine *Editors*

Immuno-proteomics

Methods and Protocols

Second Edition

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, UK

For further volumes:
http://www.springer.com/series/7651

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-bystep fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Immunoproteomics

Methods and Protocols

Second Edition

Edited by

Kelly M. Fulton and Susan M. Twine

Human Health Therapeutics Research Center, National Research Council of Canada , Ottawa, ON, Canada

Humana Press

Editors
Kelly M. Fulton
Human Health Therapeutics Research Centre
National Research Council of Canada
Ottawa, ON, Canada

Susan M. Twine
Human Health Therapeutics Research Centre
National Research Council of Canada
Ottawa, ON, Canada

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-9596-7 ISBN 978-1-4939-9597-4 (eBook)
https://doi.org/10.1007/978-1-4939-9597-4

© Springer Science+Business Media, LLC, part of Springer Nature 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover Caption: Two dimensional immunoblot for vaccine development.

This Humana imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

The mammalian immune system has evolved to constantly survey for foreign pathogens and abnormalities. The field of immunology has long sought to understand the extent of the immune response. The complexity of the immune system allows pathogens to be eliminated rapidly but also cancers to proliferate. The term "immunoproteomics" has been coined to describe the study of large sets of proteins and peptides involved in the immune system. Rapid growth of the field in the past two decades has seen the adoption of proteomics, mass spectrometry, and chip-based technologies to interrogate the immune response to infectious diseases, cancer, and autoimmunity. The resulting information has potential to be used in diagnostics, disease progression, and correlate of protection analysis, to name but a few applications.

The second edition of this book is aimed at scientists new to the field and those with years of experience in immunoproteomics. Descriptions of now classical immunoproteomics approaches are retained, many with utility as foundational techniques. This edition includes expanded sections describing the characterization of the peptides bound to major histocompatibility complexes (MHC) on the surface of cells. This complement of peptides, bound to MHC and presented to immune cells, is known as the immunopeptidome. This field has been spurred forward by advances in speed, sensitivity, and resolution of mass spectrometry. These advances are collectively advancing critical areas, such as cancer immunotherapy.

The variety of techniques presented provides not only an overview of the breadth of the field but valuable hands-on insights from specialists. Success stories will allow readers to transfer these techniques to their own laboratories, in addition to providing a reference to guide researchers.

Ottawa, ON, Canada

Kelly M. Fulton
Susan M. Twine

Contents

1 Introduction to the Immune System 1
Scott McComb, Aude Thiriot, Bassel Akache, Lakshmi Krishnan, and Felicity Stark

2 Immunoproteomics Methods and Techniques 25
Kelly M. Fulton, Isabel Baltat, and Susan M. Twine

3 Classical Immunoproteomics: Serological Proteome Analysis (SERPA) for Antigen Identification 59
Kelly M. Fulton, Anna Ananchenko, Lawrence Wolfraim, Shannon Martin, and Susan M. Twine

4 Profiling of Cytokine and Chemokine Responses Using Multiplex Bead Array Technology 79
Greg Harris and Wangxue Chen

5 Preparation of the Low Molecular Weight Serum Proteome for Mass Spectrometry Analysis 95
Timothy D. Veenstra

6 Enriching for Low-Abundance Serum Proteins Using ProteoMiner™ and Protein-Level HPLC 103
Sophie Moggridge, Kelly M. Fulton, and Susan M. Twine

7 Identification of the Antigen Content of Electroimmunoprecipitates 119
N. Helena Beyer and Niels H. H. Heegaard

8 Detection of Human Norovirus-Specific Antibodies Using the Luciferase Immunoprecipitation System (LIPS) Assay 137
Christine M. Tin and Stanislav V. Sosnovtsev

9 Intact Mass Spectrometry Analysis of Immuno-Isolated Human Therapeutic Antibodies from Serum 153
Christie E. Delaney, John F. Kelly, Wen Ding, and Arsalan S. Haqqani

10 Antigen Identification for Cell-Binding Antibodies Using Ligand-Directed Crosslinking and Biotin Transfer 167
Tammy-Lynn Tremblay and Jennifer J. Hill

11 Whole-Genome Phage Display Libraries: A Powerful Tool for Antigen Discovery 181
Elisa Beghetto and Nicola Gargano

12 Multiplexed Detection of Autoantibodies to Glycopeptides Using Microarray 199
Johannes W. Pedersen, Alexander Nøstdal, and Hans H. Wandall

13 Construction and Screening of an Antigen-Derived Peptide Library Displayed on Yeast Cell Surface for CD4+ T Cell Epitope Identification 213
Fei Wen, Mason R. Smith, and Huimin Zhao

14 Isolation of Major Histocompatibility Complex (MHC)-Associated Peptides by Immunoaffinity Purification 235
Sarah A. Penny and Stacy A. Malaker

15 Enhanced Mass Spectrometry Detection of MHC Peptides. 245
Rui Chen and Jianjun Li

16 Enrichment of Phosphorylated MHC Peptides with Immobilized Metal Affinity Chromatography and Titanium Dioxide Particles 259
Rui Chen and Jianjun Li

17 Mass Spectrometric Identification and Molecular Modeling of Glycopeptides Presented by MHC Class I and II Processing Pathways. 269
Stacy A. Malaker and Michael J. Ferracane

18 Genome-Based Bioinformatic Prediction of Major Histocompatibility (MHC). 287
Simon J. Foote

19 Enhancing Mass Spectrometry-Based MHC-I Peptide Identification Through a Targeted Database Search Approach 301
Prathyusha Konda, J. Patrick Murphy, Morten Nielsen, and Shashi Gujar

20 Prioritization of Therapeutic Targets of Inflammation Using Proteomics, Bioinformatics, and In Silico Cell-Cell Interactomics 309
Arsalan S. Haqqani and Danica B. Stanimirovic

21 Genomics-Driven Immunoproteomics: An Integrative Platform to Uncover Important Biomarkers for Human Diseases 327
Raghavendra Giri, Veneta Qendro, Pooja Rani, Carren Jepchumba, Grace Bugos, Volker Stadler, and David K. Han

22 Identification of Antigens in Immune Complexes 333
Nozomi Aibara and Kaname Ohyama

23 Analysis and Characterization of Immune Cells and Their Activation Status by Whole-Cell MALDI-TOF Mass Spectrometry 339
Richard Ouedraogo, Julien Textoris, Laurent Gorvel, Aurélie Daumas, Christian Capo, and Jean-Louis Mege

24 A Molecular Immunoproteomics Approach to Assess the Viral Antigenicity of Influenza 353
Kevin M. Downard

25 Immunoproteomic Biomarkers: From Publication to Personalized Medicine 371
Devanand M. Pinto

Index 379

Contributors

Nozomi Aibara • *Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan*

Bassel Akache • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Anna Ananchenko • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Isabel Baltat • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Elisa Beghetto • *IRBM Science Park, Rome, Italy*

N. Helena Beyer • *Department of Autoimmunology and Biomarkers, Statens Serum Institut, Copenhagen, Denmark*

Grace Bugos • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*

Christian Capo • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*

Rui Chen • *Human Health Therapeutics Research Centre, National Research Council Canada, Ottawa, ON, Canada*

Wangxue Chen • *Human Health and Therapeutics Research Center, National Research Council of Canada, Ottawa, ON, Canada*

Aurélie Daumas • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*

Christie E. Delaney • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Wen Ding • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Kevin M. Downard • *Infectious Disease Responses Laboratory, POWCS, Medicine, University of New South Wales, Sydney, NSW, Australia*

Michael J. Ferracane • *Department of Chemistry, University of Redlands, Redlands, CA, USA*

Simon J. Foote • *Human Health Therapeutics Research Centre, National Research Council Canada, Ottawa, ON, Canada*

Kelly M. Fulton • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Nicola Gargano • *Alfasigma S.p.A, Rome, Italy*

Raghavendra Giri • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*

Laurent Gorvel • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*

Shashi Gujar • *Department of Microbiology and Immunology, Dalhousie University, Halifax, NS, Canada; Department of Pathology, Dalhousie University, Halifax, NS, Canada; Department of Biology, Dalhousie University, Halifax, NS, Canada; Centre for Innovative and Collaborative Health Services Research, IWK Health Centre, Halifax, NS, Canada*

David K. Han • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*

Arsalan S. Haqqani • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Greg Harris • *Human Health and Therapeutics Research Center, National Research Council of Canada, Ottawa, ON, Canada*

Niels H. H. Heegaard • *Department of Autoimmunology and Biomarkers, Statens Serum Institut, Copenhagen, Denmark; Department of Clinical Biochemistry, University of Southern Denmark, Odense, Denmark*

Jennifer J. Hill • *Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada*

Carren Jepchumba • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*

John F. Kelly • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Prathyusha Konda • *Department of Microbiology and Immunology, Dalhousie University, Halifax, NS, Canada*

Lakshmi Krishnan • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Jianjun Li • *Human Health Therapeutics Research Centre, National Research Council Canada, Ottawa, ON, Canada*

Stacy A. Malaker • *Department of Chemistry, Stanford University, Stanford, CA, USA*

Shannon Martin • *Dynport Vaccine Company LLC, Frederick, MD, USA*

Scott McComb • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

Jean-Louis Mege • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*

Sophie Moggridge • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*

J. Patrick Murphy • *Department of Pathology, Dalhousie University, Halifax, NS, Canada*

Morten Nielsen • *Department of Bio and Health Informatics, Technical University of Denmark, Lyngby, Denmark; Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina*

Alexander Nøstdal • *Department of Cellular and Molecular Medicine, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen, Denmark*

Kaname Ohyama • *Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan*

Richard Ouedraogo • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*
Johannes W. Pedersen • *Department of Cellular and Molecular Medicine, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen, Denmark*
Sarah A. Penny • *Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK*
Devanand M. Pinto • *Human Health Therapeutics Research Centre, National Research Council Canada, Halifax, NS, Canada*
Veneta Qendro • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*
Pooja Rani • *Department of Cell Biology, Center for Vascular Biology, University Connecticut School of Medicine, Farmington, CT, USA*
Mason R. Smith • *Department of Chemical Engineering, University of Michigan, Ann Arbor, MI, USA*
Stanislav V. Sosnovtsev • *Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA*
Volker Stadler • *PEPperPrint, GmBH, Heidelberg, Germany*
Danica B. Stanimirovic • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*
Felicity Stark • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*
Julien Textoris • *Aix Marseille Université, Unité de Recherche sur les Maladies Infectieuses Tropicales et Emergentes, CNRS UMR 7278, IRD 198, INSERM U1095, Marseille, France*
Aude Thiriot • *Division of Immunology, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA*
Christine M. Tin • *Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA*
Tammy-Lynn Tremblay • *Human Health Therapeutics, National Research Council Canada, Ottawa, ON, Canada*
Susan M. Twine • *Human Health Therapeutics Research Centre, National Research Council of Canada, Ottawa, ON, Canada*
Timothy D. Veenstra • *Department of Applied Sciences, Maranatha Baptist University, Watertown, WI, USA*
Hans H. Wandall • *Department of Cellular and Molecular Medicine, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen, Denmark*
Fei Wen • *Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Department of Chemical Engineering, University of Michigan, Ann Arbor, MI, USA*
Lawrence Wolfraim • *Dynport Vaccine Company LLC, Frederick, MD, USA*
Huimin Zhao • *Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Department of Chemistry, Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA*



Chapter 1

Introduction to the Immune System

Scott McComb, Aude Thiriot, Bassel Akache, Lakshmi Krishnan, and Felicity Stark

Abstract

The immune system in a broad sense is a mechanism that allows a living organism to discriminate between "self" and "nonself." Examples of immune systems occur in multicellular organisms as simple and ancient as sea sponges. In fact, complex multicellular life would be impossible without the ability to exclude external life from the internal environment. This introduction to the immune system will explore the cell types and soluble factors involved in immune reactions, as well as their location in the body during development and maintenance. Additionally, a description of the immunological events during an innate and adaptive immune reaction to an infection will be discussed, as well as a brief introduction to autoimmunity, cancer immunity, vaccines, and immunotherapies.

Key words Immune system, Immunity, Vaccines, Adjuvants, Checkpoint, CAR-T, Adaptive immunity, Innate immunity, Inflammation

1 Introduction

Early immune system discoveries were largely fuelled by a desire to prevent the spread of disease and develop better treatments for the sick (Fig. 1). As far back as the eighteenth century, microbiologists sought to inoculate healthy people against diseases. In fact, vaccines were created to combat illness before anyone could prove that microbes caused illness, or that immune cells could kill microbes. Over 100 years prior to Koch's postulates in 1890, which definitively identified microbes as the causative agent of disease, Edward Jenner had made a crude vaccine from the pus of cowpox lesions to successfully immunize people against smallpox [1].

As microbiologists were uncovering the properties of bacteria, it was commonly believed that white blood cells aided the spread of pathogens by transporting them throughout the host. In 1882, Eli Metchnikoff recognized that white blood cells were instead destroying the engulfed pathogen [2]. The word phagocyte, from

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_1, © Springer Science+Business Media, LLC, part of Springer Nature 2019



Fig. 1 The term "immune" is derived from the Latin immunis, literally meaning to be exempt from the public service required of lower classes. While the Roman empire may have bestowed immunity from taxation upon the chosen few, there was no special exemption from the scourge of virulent disease

the Greek words "phagein," to eat, and "cyte" cell, was used to describe this cellular action. Metchnikoff had identified an important process which is part of a greater network of first line of defense mechanisms known as *innate immunity*. Innate immunity is the evolutionarily older arm of the immune system, composed of the barriers (skin), small molecules (complement), and cells such as macrophages and dendritic cells. The innate immune system is so-called because it provides protection from pathogens without the need for preconditioning from the environment. In other words, when the innate immune system encounters a pathogen, it will react immediately to kill or to remove it from the host.

In addition to Metchnikoff's discovery of cellular immunity, other researchers were examining the ability of bodily fluids (humors) to provide protection against disease. In 1890, Emil von Behring and Shibasaburō Kitasato discovered antibodies when they identified acellular components of the blood that conferred immunity when transferred from one animal to another [3]. Antibodies, along with cytokines and complement, are the components of *humoral immunity*. Interestingly, the discovery of antibodies

sparked heated debates that divided scientists about the importance each type of immunity played in overall host immunity. This divide was bridged in 1903 when scientists Almroth Wright and Steward Douglas proved that humoral responses aided the cellular immune response suggesting that both cellular and humoral immune responses played important roles. They observed that antibodies and complement enhanced the phagocytosis of bacteria by binding to the bacteria, an event termed opsonization [4].

In contrast to humoral immunity, the cellular fraction of the immune system is more commonly known as *cell-mediated immunity*. The ability of the immune system to generate a specific cellular response to a pathogen became known as *adaptive immunity*. How the immune system was able to produce such highly specific immune receptors remained a long-standing mystery in immunology until the mid-twentieth century. In 1965, Dreyer and Bennet published a speculative paper suggesting that DNA recombination of immune genes could generate immune diversity [5]. Throughout the next decade, the work of many researchers identified that the V(D)J genetic elements of lymphocyte receptors could be shuffled randomly. Through this rearrangement, a library of lymphocytes is created each with a unique gene for its immune receptor. The result is a large pool of highly diverse lymphocytes each capable of binding a unique pathogen associated molecule. The discovery of lymphocyte receptor specificity could explain Emil von Behring and Shibasaburō Kitasato's earlier findings that sera from vaccinated animals could provide protection to other animals challenged with the same pathogen. The antibodies present in the sera were actually soluble immune receptor which is a product of a clone of B lymphocytes.

While the practical application of immunizations may have outpaced our understanding of the immune system in the early years of modern medicine, we now have insight into the machinery behind immunity. This knowledge has translated into better tools to both monitor and manage immune responses and ultimately improve patient outcomes. In this chapter we will further delve into the mechanisms that allow our bodies to recognize, respond to, and remember pathogens which challenge the body.

2 Components of the Immune System

2.1 Organs of the Immune System

While immune cells can be found throughout the body, the immune system also has a specialized network of immune organs (Fig. 2). The organization of the immune system within immune organs allows for a regulated immune response capable of rapidly producing a large number of cells that can halt a spreading infection. From these immune reservoirs, immune cells and molecules

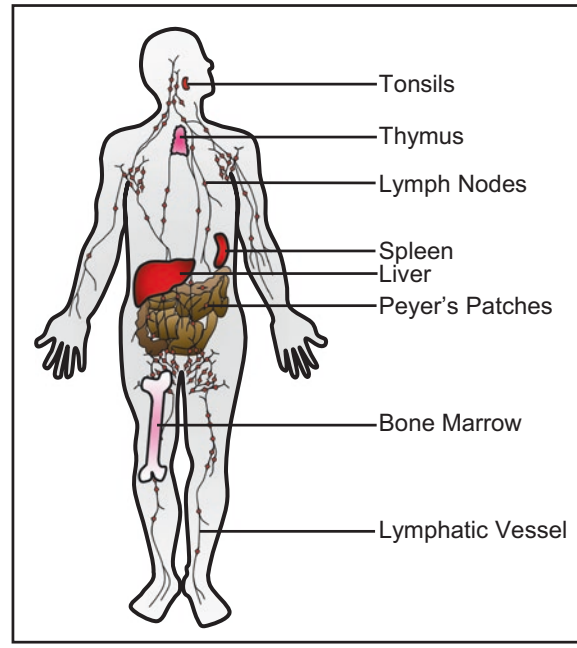


Fig. 2 While immune cells are located throughout the body, the immune system relies on specialized organs to generate immune cells, detect pathogens, and initiate immune responses. Many immune cells arise from precursor cells located in the bone marrow. In the case of T cells, they undergo genetic recombination to form a TCR in the thymus. T and B cells with mature receptors then migrate through lymphatic vessels to lymph nodes where they await activation signals. Large organ systems also have specialized immune sites that harbor immune cells such as the spleen for the circulatory system or Peyer's patches for the gut

can be released to penetrate almost any tissue throughout the body.

While all cells of the blood originate from the same precursor hematopoietic stem cells in the bone marrow, their sites of maturation and residence differ [6]. Many different types of innate immune cells typically arise from the bone marrow and take residence in the blood and tissues (discussed further below). For cells of the acquired immune system, T and B cells, they recombine their immune receptors in the thymus and bone marrow, respectively; these are known as *primary immune organs*. After maturation in these primary sites, T and B cells will reside in the lymphatic tissue, and as such they are often referred to as lymphocytes. These sites of lymphocyte residence are known as the *secondary immune organs*, and they include the lymph nodes, spleen, Peyer's patches, the appendix, tonsils, adenoids, and other mucosal-associated lymphoid tissue (MALT).

The transportation of immune cells throughout the body occurs both in blood and lymph. Lymph is a clear fluid made up of proteins of the interstitial fluid of bodily tissues; it accumulates in

lymph capillaries which are thin-walled vessels dispersed throughout all tissues (except the central nervous system (CNS)). Lymph capillaries differ from blood capillaries in that they allow only a one way passage of lymph into and not out of them. Capillaries lead to the afferent lymph vessels that lead into the lymph node. Unlike the circulatory system of blood, the lymphatic system is not a closed circulatory system, and all lymph is moved from lymph nodes to efferent lymphatic vessels toward the heart, unidirectionally. Lymphoid fluid moves through lymph vessels by the passive actions of skeletal muscle contractions and eventually is returned to the blood at the junction of the internal jugular and subclavian veins at the base of the right side of the neck [7]. Using the bloodstream and lymph, immune cells and pathogen debris can enter lymphoid organs to generate immune responses.

While some innate immune cells are tasked with locating pathogens throughout the body and killing them by various means, other innate immune cells such as dendritic cells have a more complex task to fulfil. Once a dendritic cell has engulfed pathogen, it will migrate toward a lymphoid site within the blood or lymph. The dendritic cell will migrate through the lymphoid tissue with a goal of encountering a lymphocyte (T or B cell) and that initiates an adaptive immune response which typically occurs within a few days (Fig. 3). Thus it can be envisioned that the lymphoid compartment is the meeting ground between innate and adaptive immune cells that initiate the activation of the adaptive immune response.

2.2 Cells of the Immune System

The cells of the immune system have been classified into two general groups; these are innate or adaptive immunity (Fig. 4). Innate immune cells react quickly, whereas adaptive immune cells have a delayed response that can take days to fully develop but go on to form immunological memory.

The quick responding innate immune cell types include granulocytes (polymorphonuclear cells), mast cells, macrophages, and dendritic cells. *Mast cells* are best known for their ability to rapidly release granules of histamine and heparin in response to an infection. This rapid response can be important in initiating inflammation and wound healing but is also involved in allergic responses. *Granulocytes* encompass a group of three cell types differentiated by the contents of their granules: neutrophils, basophils, and eosinophils. All three are relatively short-lived (~5 days) but are important early responders to parasites, extracellular bacteria, and tumors. The early arrival of granulocytes during infection induces acute wound inflammation and dilation of the surrounding blood vessels that allow for the rapid influx of other immune cells. *Neutrophils* are of particular importance as they make up about half of the circulating white blood cells in humans and have a keen ability to phagocytose and destroy invading microbes. In addition to

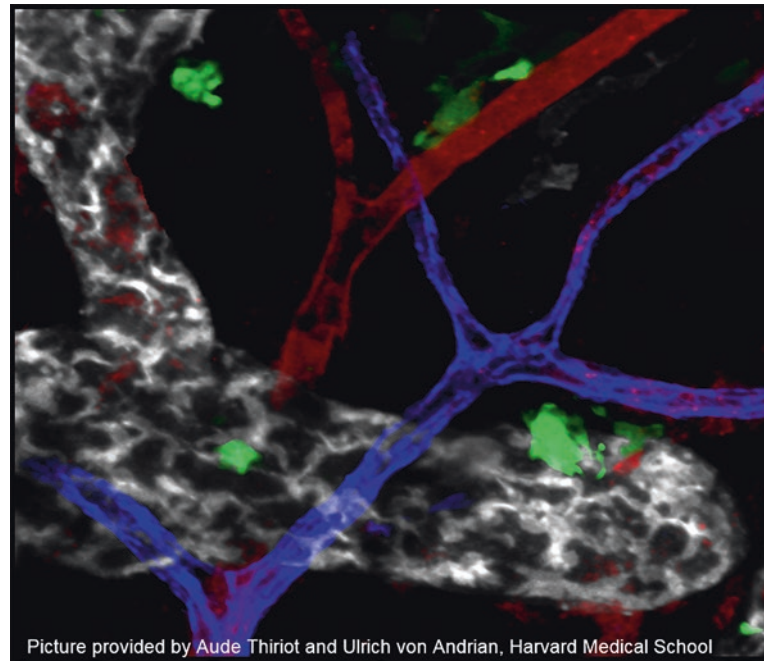


Fig. 3 Immune cells are waiting and watching for invading pathogens throughout the body. Shown here, dendritic cells (CD11c—green) within the skin sit near blood vessels (CD31—red) and venules (DARC—blue). If these sentinel dendritic cells detect a pathogen, they will trap the invader and use a network of lymphatic vessels (LYVE-1—white) to move to a lymph node where an immune response can be activated

secretion of inflammatory mediators such as cytokines and soluble antimicrobial peptides, neutrophils have recently been found to also undergo a suicidal extrusion of neutrophil extracellular traps (NETs). NETs are primarily composed of DNA and provide a physical barrier by which pathogen is trapped and prevented from spreading [8].

Similar to neutrophils, macrophages are adept phagocytic cells, capable of ingesting and destroying invading microbes. Unlike neutrophils, which are blood-resident and short-lived, macrophages take up residence in all tissues of the body and are relatively long-lived. Macrophages are also well known to induce inflammation by producing cytokines and chemokines that both draw in and activate other immune cells to the site of infection.

Dendritic cells (DCs) are also capable of internalizing and destroying invading microbes through phagocytosis; however they are better known for their important role of activating the cells of the adaptive immune system. Both DCs and macrophages are often called antigen-presenting cells (APCs) due to their efficient ability to internalize pathogen and present pathogen peptides on their cell surface. The term *antigen* refers to a molecule that can evoke a

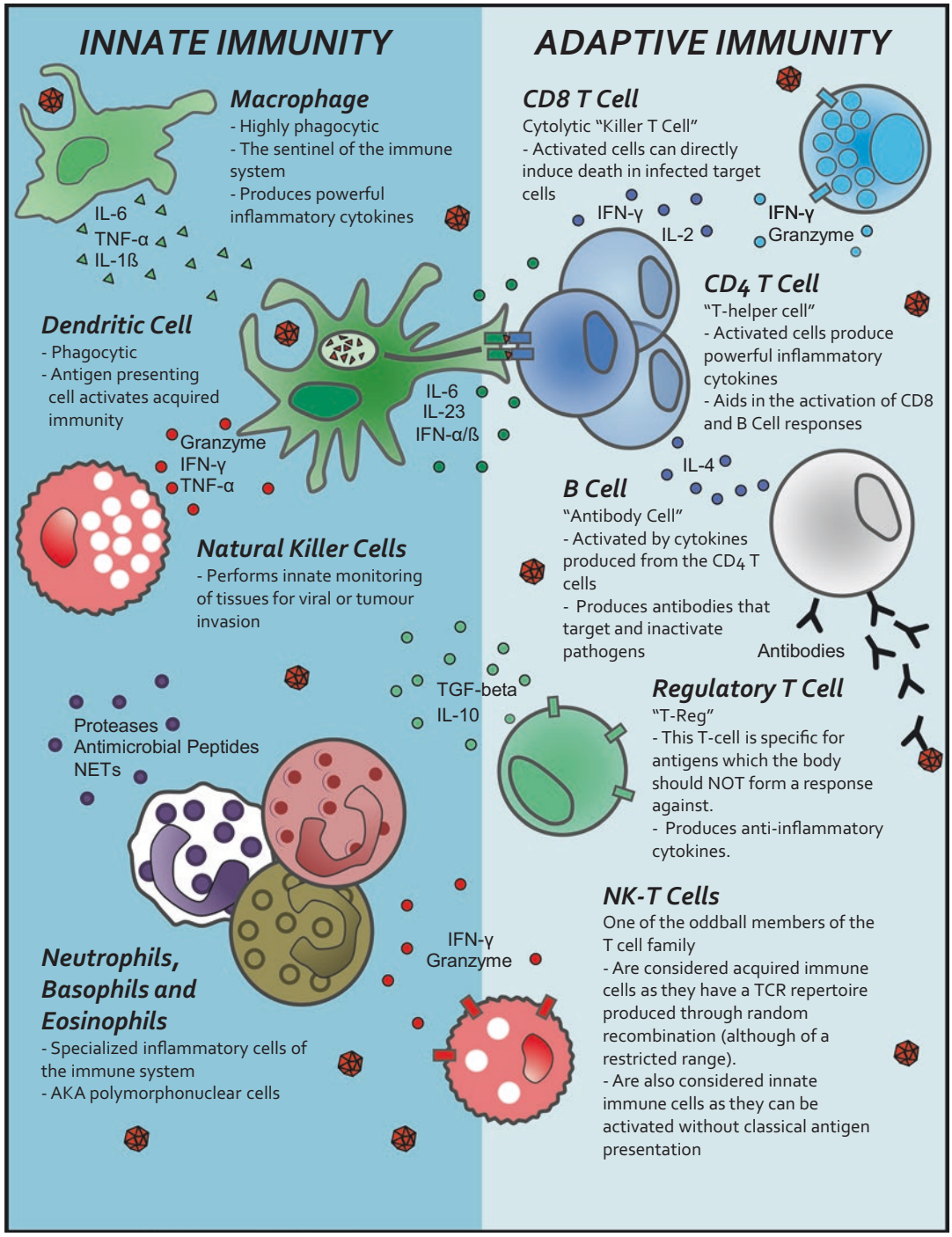


Fig. 4 Common cells of the immune system

specific adaptive immune response, and in this case, the antigen is the pathogen peptide fragments.

Cells, such as DCs and macrophages, which are capable of antigen presentation, can be subclassified as either *professional or nonprofessional APCs.* Both classes of cells process and present pathogen peptide on a receptor called the major histocompatibility complex class II (MHC-II). More specifically, upon APC internalization of the microbe or debris, peptide fragments are generated by proteasomal degradation. The peptide fragments then become bound to the MHC-II receptor and are shuttled to the surface of the cell. Within a lymphoid site, APCs will come into contact with and activate lymphocytes. This occurs via the engagement of the MHC-II with the T-cell receptor (TCR); this interaction is often referred to as "signal 1," but it alone is not enough to activate naive T cells. *Professional APCs* such as macrophages, DCs, and B cells also express cell surface co-stimulatory molecules such as CD80, CD86, and CD40, which can provide "signal 2" to a naive lymphocyte. Importantly, a lymphocyte will become activated only if it receives both signals 1 and 2; otherwise it can go into a stasis known as anergy. A *nonprofessional APC*, such as a fibroblast, thymic epithelial cell, or vascular endothelial cell, only expresses MHC-II in the presence of certain cytokines and does not express co-stimulatory molecules. Thus they are unable to activate naive T cells but can play a role in reactivating memory lymphocytes.

T cells are generally classified into two groups expressing either cell surface CD4 or CD8 receptors. CD4 and CD8 play an important role in the formation of the immune synapse between the TCR and MHC of lymphocytes and target cells, respectively. *CD8 T cells* are most commonly known as cytotoxic T lymphocytes (CTL) because once they strongly engage a target cell they secrete cytotoxic granules and perforin into the immune synapse that penetrate the target cell and induce apoptosis. *CD4 T cells* are commonly referred to as helper T cells because they play an important role in contributing to the cytokine response that drives either cell-mediated immunity by macrophages and CD8 T cells or humoral immunity mediated by B cells. When CD4 T cells are activated in the presence of IL-12 and IFNγ, they become of the Th1 phenotype and secrete interferon-γ (IFNγ) and lymphotoxin-alpha (LT-α) into the environment which induces inflammation and supports the function of macrophages and CD8 T cells to kill pathogens. However, when CD4 T cells are activated in the presence IL-4, they become of the Th2 phenotype which supports of the action of B cells to produce antibodies.

B cells provide surveillance to the body for signs of infection by circulating in the blood and lymph. Their antigen recognition receptor known as the B-cell receptor (BCR) is actually a cell-

bound antibody. Once the B cell becomes activated in the presence of their target antigen, it becomes a plasma cell and begins to produce and secrete large amounts of antibody that can bind to the target protein and neutralize it (Fig. 5). B cells are also classified into two major populations named B1 and B2 lymphocytes, according to their cell surface markers, anatomical location, and immunological function. B1 cells are enriched in pleural and peritoneal cavities and are the main producers of natural antibodies, in particular IgM. B2 cells consist of two subsets, the marginal zone B (MZ B) cells and the conventional follicular B (FO B) cells. MZ B cells reside in the marginal zone of the spleen and are involved in innate immune response. FO B resides in the follicular zone in the spleen and is present in other lymphoid organs such as lymph node. FO B cells are involved in the adaptive immune response.

While the lymphocytes mentioned thus far predominantly contribute to the adaptive immune response, there exists another lymphocyte termed *natural killer cells* (NK) that contributes to innate immune responses. The NK lymphocyte performs directed cytolytic killing similar to CD8 T cells by secreting cytotoxic granules into the immune synapse. Their specialized receptors (NKG2D, KIR, etc.) scan neighboring cells for signs of infection by recognizing the absence of cell surface MHC. A mechanism by which pathogens often evade the immune system is to prevent the expression and translocation of MHC-peptide molecules to the cell surface. NK cells detect this concealment and kill any cells that do not express a certain amount of MHC molecule. Since NK cells do

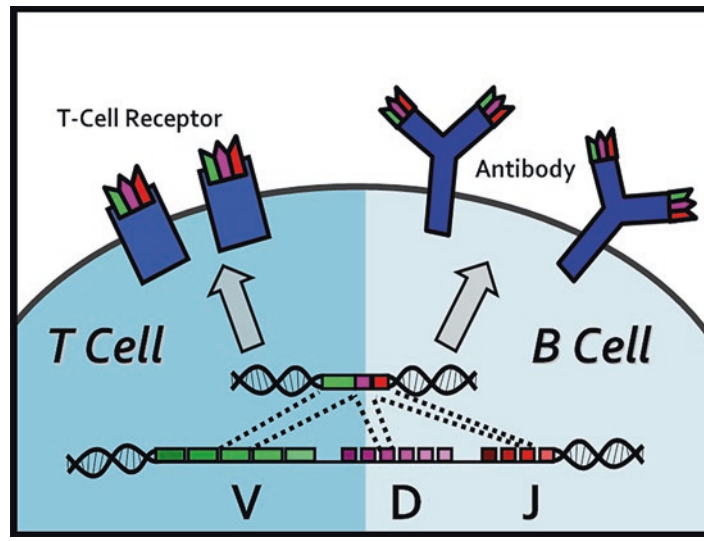


Fig. 5 Somatic DNA recombination generates lymphocyte receptor diversity

not require activation to kill, target cells such as tumor or virally infected cells can be killed within 3 days compared to CD8 T cells that typically require 5–7 days to start killing. NK cells are not thought to express genetically rearranged antigen receptors. However, recent findings have revealed that a subset of liver-resident NK cells can mediate adaptive immune responses. Indeed, this hepatic NK cell population can acquire long-lived and highly specific memory of a variety of viral and hapten-based antigens [9]. Unravelling the molecular mechanisms by which these NK cells have specificity to antigens is a hot topic in immunology since they do not undergo the V(D)J gene rearrangement that give B and T cells antigen specificity. However, it has been shown that the liver is home to antigen-specific memory NK cells that have at one point expressed the RAG enzymes normally used in V(D)J gene rearrangement and that this correlates with improved functionality and longevity of antigen-specific NK cell responses [9–12]. Exploring this newly discovered arm of adaptive immunity could open many doors in the development of new vaccines.

A number of other immune cell subsets perform important roles in an immune response; these include *regulatory T cells* and *suppressor macrophages* that play a role in suppressing allergy and autoimmune diseases. There also exists a unique cell subset called the *NK-T cell*, which has both cell surface receptors belonging to NK cells and T cells and is thought to be involved in the suppression of cell-mediated autoimmune responses [13]. The *Th-17* subset of CD4 T cells has also been of interest to researchers as its production of IL-17 has been implicated as a causative agent of autoimmunity; however this same cell has been shown to play a protective role in pulmonary infections [14]. While immune cells have been generally classified into groups according to their most prominent roles in an immune response, the constant discovery of novel secreted small molecules and receptors has complicated the categorization process but has also opened new avenues of research to identify new cellular and molecular targets to harness in the prevention of disease.

2.3 Adaptive Immune Receptors

2.3.1 Generating the Adaptive Immune Receptor Library

Adaptive immunity is one of the most powerful elements of the immune system because it selects for the most appropriate immune receptor to target the infecting pathogen. Through this selection process, the adaptive immune system can also "remember" pathogens by maintaining pathogen-specific memory cells. Unlike other cells within the body, T and B cells undergo changes at the DNA level during development. Through a mechanism of somatic recombination, gene elements are cut and pasted resulting in. random recombination. This creates a staggering diversity of T and B cell receptors.

B cells for example can potentially produce antibodies specific for up to ~10^{18} unique targets, whereas T cells can produce ~10^{13} different receptors [4]. While new T and B cells are produced throughout an organism's lifetime, once a T cell undergoes somatic

DNA recombination it maintains this receptor for the remainder of its life. In contrast, activated B cells will undergo further genetic alterations through somatic hypermutation, gene conversion, and class switching. This genetic change further diversifies the specific B cell receptor which is already capable of binding a target. Through successive rounds of mutation and selection, the antibody affinity can be improved severalfold.

2.4 Cytokines

Cytokines encompass a wide variety of signaling peptides, proteins, and glycoproteins that are used in intercellular communication (Table 1). The categorization of cytokines into tidy groups has been challenging since new properties of existing cytokines are constantly being uncovered. Historically, cytokines were named either by their cellular origin or their cellular target; i.e., interleukins are produced by leukocytes, and tumor necrosis factor (TNF) inhibits tumorigenesis. Since the discovery that some previously identified interleukins were also produced by other cell types, the term interleukin has shifted to define a broader set of cytokines and is often used to label newly discovered cytokines. Cytokines have also been classified according to their various functions and can be found within the following groups. *Interleukins* are by far the broadest grouping of cytokines as the term interleukin is derived from the Latin "occurring between leukocytes." That definition can explain the actions of almost all cytokines and is likely why newly discovered cytokines are given that designation. *Chemokines* are small (8–10 kDa) and are so named due to their ability to attract other cells. *Lymphokines* include those cytokines produced by lymphocytes; thus, some cytokines such as IL-2, IL-6, and IL-10 can be both a lymphokine and an interleukin. *Interferons* (IFN) are a subset of cytokines that share a common ability to combat cancer and viral infection. They are glycoproteins and segregated into two groups, type I (IFNα and IFNβ) and type II (IFNγ). The original *tumor necrosis factor (TNF) family* members include TNFα and TNFβ; the latter is now referred to as lymphotoxin alpha (LTα). Originally named for their ability to mediate tumor cytotoxicity, TNF family members (CD40L, FasL, etc.) are now identified based on sequence and structural similarities and mediate a variety of functions from the regulation of cell differentiation to cell survival. Table 1 summarizes common cytokines, their cellular source, function, and molecular weight.

The use of cytokines to combat cancer and autoimmunity has been heavily studied as many cytokines have well-defined functions and can be synthetically engineered relatively easily. Cytokines have been approved for use in cancer therapy as they have been shown to support both innate and adaptive immune responses to combat cancer; these include IFNα, IL-2, and GM-CSF [15]. However, limited success has been observed in general with cancer cytokine therapy due to high toxicity and failure to reach an effective concentration in the tumor. Cytokines are currently being investi-

Table 1
Common cytokines, their cellular source, function and size [11]

Cytokine	Produced by	Actions	Size (kDa)[a]
IL-2	T cells	T cell proliferation	16, monomer
IL-4	T cells, mast cells	B cell activation and Th2 cell differentiation	16 and 18[b], monomer
IL-6	T cells, macrophages, endothelial cells	T and B cell growth and differentiation	24, monomer
IL-7	Thymic stromal cells, lymphatic endothelial cells	Homeostatic proliferation of naive T cells and memory CD8 T cells	15 and 20[b], monomer and dimers
IL-10	Monocytes, Th2 T cells and Treg	Macrophage suppression and inhibition of Th1 cell differentiation	21, homodimer
IL-12	Macrophages and dendritic cells	NK cell activation Th1 cell differentiation	37 and 25[b], heterodimer
IL-15	Mononuclear phagocytes	Stimulation of T and NK cell growth and promotion of memory CD8 T cell survival	15 and 18[b], monomer
IL-17	T cells and macrophages	Induction of epithelial, endothelial and fibroblasts to produce proinflammatory cytokines	18, homodimer
IL-21	Th2 and Th17 T cells	Induction of T, B and NK cell proliferation	17 and 18[b], monomer
CD40 L (CD154)	T cells and mast cells	Activation of B cells and class switching	29, monomer, dimer and trimer
Lymphotoxin (LT, TNFβ)	Th1 and CD8 T cells	Activation of macrophages and neutrophils and inhibition of T cells and tumours	22, 33[c], trimer
Interferon-α	Leukocytes and dendritic cells	Antiviral	19, monomer
Interferon-β	Fibroblasts	Antiviral	19, monomer
Interferon-γ	Th1, CD8 T cells and NK cells	Activation of macrophages and NK cells. Inhibition of Th2 differentiation	19, 25[b], dimer and tetramer
G-CSF	Fibroblasts and monocytes	Promotion of neutrophil development	22, monomer

(continued)

Table 1
(continued)

Cytokine	Produced by	Actions	Size (kDa)[a]
GM-CSF	T cells and macrophages	Promotion of dendritic cell differentiation. Activation of macrophages	16, 35[c] monomer
M-CSF	T cells, bone marrow stromal cells, osteoblasts	Promotion of macrophage development	α: 29 β:60 Λ: 60 Homodimer
TGF-β	CD4 T cells and Tregs	Inhibition of T cell growth but promotion of survival. Inhibition of macrophage activation	1: 44 2: 48, 50[b] 3: 47 25[d], homo and heterodimer
TNF-α	T cells	Inhibition of tumorigenesis	26, 17[d], trimer

[a]Unless otherwise indicated sizes are for the unprocessed precursor protein.
[b]Isoform
[c]Glycosylated
[d]Active, cleaved protein

gated in combinatorial treatments against cancer including combinations with checkpoint inhibitors, viral therapy, and monoclonal anticancer antibodies, among others [16]. Cytokines used to treat autoimmune diseases are usually of an anti-inflammatory nature and include IFNβ for multiple sclerosis [17] and IL-10 for psoriasis [18] and Crohn's disease [19]. Due to the potent nature of many cytokines and reports of toxicity related with intravenous delivery, new approaches for cytokine therapy include the targeted delivery of cytokines to tissues using antibodies or the slow release of cytokines encapsulated inside liposomes.

More recently, with advancing technology that increases detection sensitivity and throughput of biological samples, the measurement of cytokine concentration in serum and tissues has been considered as a diagnostic criterion. Clinical researchers have found unique cytokine profiles in disease states such as heart failure [20] and liver toxicity [21]. While the presence of a cytokine cannot point to a specific disease as can be done with a single antibody, the cytokine milieu can provide a fingerprint of a person's state of health that can arguably be used to diagnose disease.

3 Inflammation and the Innate Response

In humans, the skin and mucosal membranes comprise the first line of defense barriers against a pathogen. When disrupted by infection or lesion, the innate immune mechanisms in the skin begin to activate almost immediately.

One of the fastest acting mechanisms of the immune system is the *complement* reaction. When blood-resident inactive complement proteins encounter a pathogen, they will rapidly bind to it, either through direct interaction or in cooperation with a pathogen-specific antibody. Upon binding to pathogen, a series of proteolytic cleavages will activate the complement proteins; this causes the formation of large multimeric complexes that in turn disrupt bacterial membranes, killing an invading pathogen directly.

Macrophages, DCs, and neutrophils form the next wave of the immune response to follow the acellular mechanisms such as complement. These innate immune cells are able to sense and home to danger signals associated with damage and/or infection. In addition, they will continuously sample their local microenvironments through phagocytosis. Pathogen-associated molecular patterns (PAMPS), such as lipopolysaccharide (LPS), flagellin, single-stranded RNA, and unmethylated CpG, among others, bind to cell surface Toll-like receptors (TLRs) on innate immune cells, causing the cells to become activated. Activated innate immune cells will begin production of pro-inflammatory cytokines and interferons, further amplifying the inflammatory response. Inflammatory cytokines may also have the effect of interrupting the spread of a pathogen; for example, IFNα and IFNβ can make nearby cells much more resistant to infection with viral pathogen.

In addition to binding PAMPs in order to recognize and initiate responses against pathogens, a new class of immune activating molecules known as damage-associated molecular patterns (DAMPs) has recently been defined. This group of immune stimulators is mostly composed of intracellular molecules which, once released into the extracellular environment, can activate innate immune cells and induce inflammation. The immune-activating function of DAMPs may be particularly important in the case of cancer, as these altered cells lack any foreign molecular patterns which would classically be necessary to induce an immune response. The cellular damage incurred by unregulated growth of cancer cells can drive the recruitment of innate immune cells such as NK cells which can target and eliminate cancer.

In many cases the infecting pathogen is killed by the swift actions of the innate immune response described above. However in some cases more action is required. While some innate immune cells are involved in the direct pathogen attack, macrophages and DCs will also transport pathogen and/or pathogen debris into a lymph node. Naïve T and B cells continuously circulate from the blood to the lymph nodes via specialized microvessels, the high endothelial venules (HEVs). HEVs express a variety of adhesion molecules such as selectins, integrins, members of the immunoglobulin superfamily, and some mucin-like molecules that mediate the interactions with lymphocytes which allow them to migrate into the lymph node (Fig. 6). Antigen-specific lymphocytes in

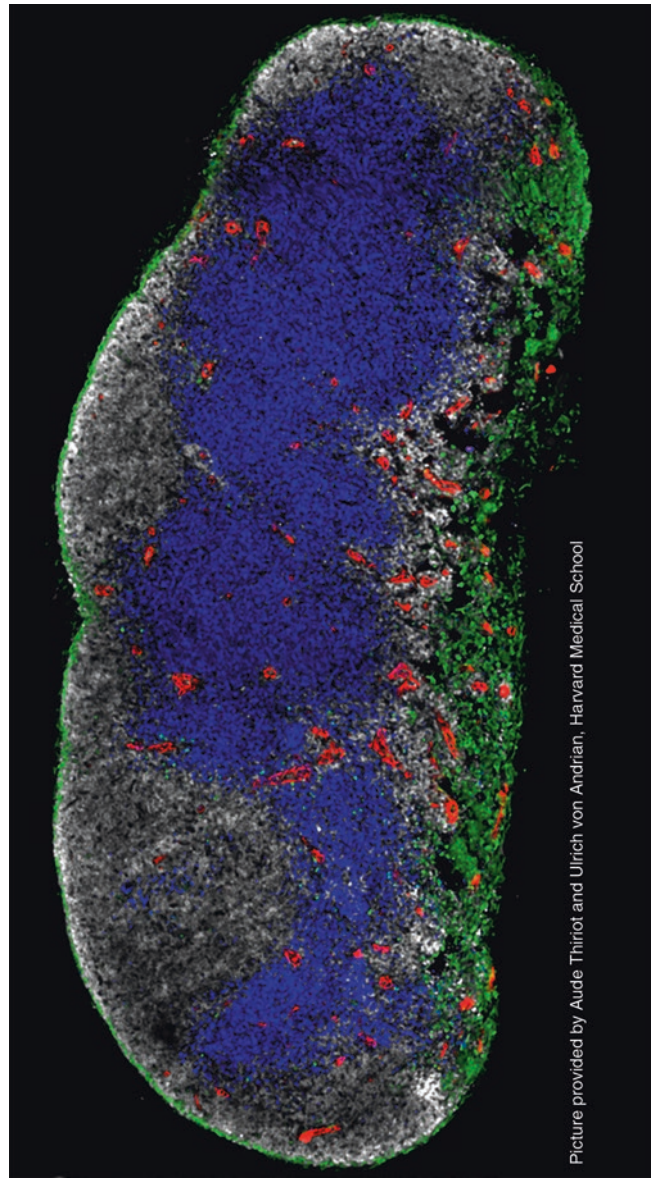


Fig. 6 The immune system relies on specialized organs, such as the lymph node shown above, where innate and adaptive immune cells come together. Macrophages (CD169—green) and dendritic cells phagocytose pathogens and digest them into small pieces. Dendritic cells then present small peptide fragments of the pathogens on their cell surface. T cells (TCRβ—blue) are in turn activated by the antigen-presenting dendritic cells and begin to proliferate. Additionally, B cells (CD19—white) can be activated by T cells and start to produce pathogen-specific antibody. Blood vessels are shown in red

lymphoid tissues are activated by APCs, which marks the beginning of a slower but much more targeted adaptive immune response.

4 Immune Activation and the Acquired Response

In contrast to the broad specificity of innate immune cells toward pathogens, adaptive immune cells are antigen specific which means that of the billion or so T and B cells in the human body only a handful (~10–100) are specific for any given pathogen protein. To streamline the interaction of an APC with its antigen-specific T cell, T cells will circulate continuously throughout the lymphoid organs increasing their opportunity to eventually encounter their target.

As described earlier, APCs use a degradative pathway to break down phagocytosed debris into peptide pieces. These peptides become bound to MHC molecules and are then shuttled to the cell surface where they can encounter the TCR of adjacent T cells. There are two types of MHC receptors: MHC class I and MHC class II. Generally, proteins derived from an intracellular pathogen (viruses and bacteria) will bind to MHC-I, while those proteins derived from the extracellular environment (bacteria and parasites) will be bound to MHC-II. While all nucleated cells express MHC class I, only a handful of cells are capable of expressing MHC class II. Of all the cells capable of expressing MHC molecules, DCs are considered the most adept at activating and inducing the proliferation of naive T and B cells because they possess co-stimulatory molecules and migrate more efficiently through lymphoid tissue to encounter a T cell.

The vast majority of TCR/MHC-peptide interactions will be of low affinity, last for a short period of time, and will not lead to T-cell activation. However when a T cell encounters an APC and its TCR has a high affinity for the MHC-peptide complex, the cell-cell interaction is reinforced with adhesion receptor/ligand interactions (co-receptors), and the cells remain in contact for a longer period of time. It is important to note that the APC will express these co-receptors only if it has received adequate activation signals from inflammatory cytokines and/or it directly receives activation signals through stimulation of its own TLRs. In addition to this, inflammatory cytokines released from the APC or other inflammatory cells can have a direct impact on the T cells. Thus, a T cell can only become activated if it has the TCR specific for the pathogen peptide in addition to the cytokine signals derived from inflammation associated with an infection (Fig. 7).

The point of interaction between the T cell and the antigen-presenting cell is known as the immunological synapse. Upon T-cell activation, many biochemical signaling complexes cause the

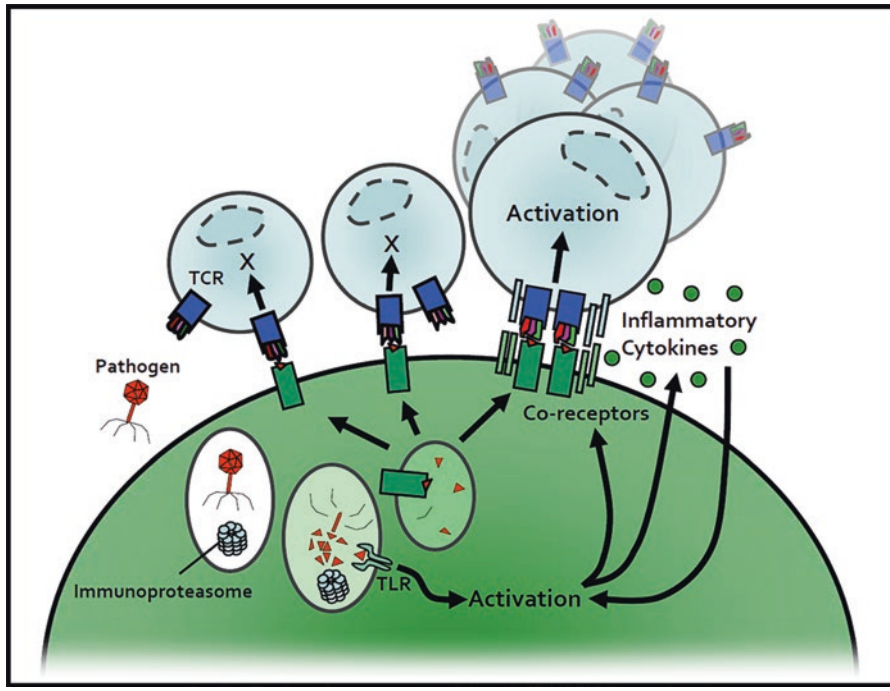


Fig. 7 An adaptive immune response relies on the sequential activation of different types of immune cells. Antigen-presenting cells (green) become activated when they encounter a pathogen. Pathogen-derived peptides bind to MHC molecules on antigen-presenting cells and engage the T-cell receptor on T cells. Only strong receptor ligand interactions will cause the activation of T cells

rapid clonal proliferation of the T cell, as well as the secretion of IL-2 by T cells that support their proliferation. During this rapid expansion, activated T cells undergo differentiation to become "active" cells capable of exerting either direct cytolytic killing or rapid cytokine production within 7 days.

Depending on which subtype the T cells belong to, different "active" functions are performed. In the case of *CD8 T cells*, they move into the periphery where they scan for their cognate ligand, namely, the MHC-I receptor bound to peptide. MHC-I receptor is present on all nucleated cells; therefore almost any cell type is subject to CD8 T-cell scanning. If the activated CD8 T cell binds tightly to a target cell, it will release cytolytic granules into the cell synapse that penetrate the target cell causing its death. This allows for the clearance of intracellular pathogens by killing the infected cells. In the case of "helper" *CD4 T cells*, they can have divergent effects depending on the type of activation they received. If activated in the presence of IL-12 and IFNγ, CD4 T cells will aid in the clearance of a pathogen via the activation of CD8 T cells. This type of cell-specific response is known as a type-I T-helper (*Th1*) response and is usually caused by intracellular infections caused by viruses, bacteria, and protozoan. When CD4 T cells are activated

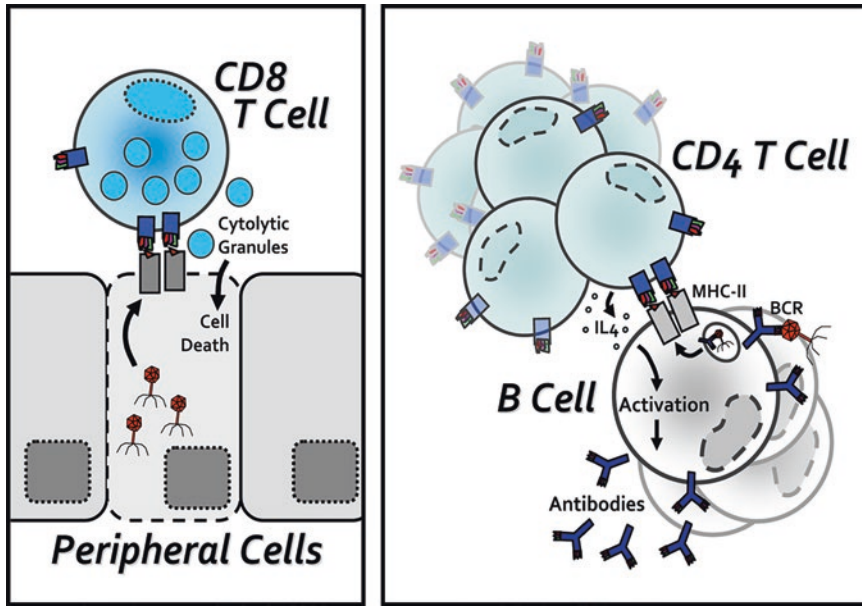


Fig. 8 T-cell subsets have different functions. Activated CD8 T cells (left) move into peripheral organs and scan the MHC-I-peptide complexes presented by all cells. If the CD8 cell detects its target antigen, then it will form a strong bond with the cell and transmit a death signal which kills the infected cell. In the case of activated Th2 CD4 T cells (right), these cells will bind to MHC-II-antigen complexes presented by B cells in the lymph node. If the CD4 T cell detects its target antigen, then it will bind strongly and transmit an activation signal to the B cell. Activated B cells will then begin to produce antibody

in the presence of IL-4, they cause the activation of a B-cell-mediated antibody response. This type of CD4 T-cell response is known as a type-II T-helper (*Th2*) response (Fig. 8).

Naive *B cells* in the lymph node will display a cell-bound antibody on their cell surface, known as the BCR. By displaying their BCR to the extracellular environment, the cells are constantly scanning, waiting for something to bind its receptor. Once the BCR is engaged, the receptor/ligand complex will then be internalized by the B cell. In some cases an entire pathogen is internalized with the receptor complex. The B cell then digests antigens from the pathogen and presents the antigenic peptides upon an MHC-II receptor. If the B cell then encounters an activated CD4 T cell with its cognate TCR, it will form a synapse with the CD4 T cell. Through the release of several cytokines (such as IL-4), the B cell will become activated. Following activation, B cells will undergo massive expansion and differentiation, similar to what occurs in T-cell activation. Unlike T cells however, B cells will change the way in which their specific receptor is expressed. Rather than displaying their BCR on the cell surface, the B cell will release its receptor as an antibody. At this point, the B cell is known as a *plasma cell* and will produce a large quantity of antibody which will

Table 2
The various forms of antibodies and their functions

Antibody	Function(s)	Location(s)
IgM	• Naïve BCR • Early antibody production • Activates the complement system	Membrane bound, released as a pentamer
IgD	• Naïve BCR	Membrane bound
IgG	• Directly neutralize target proteins • Mark targets for phagocytosis • Mark targets for neutrophil degranulation • Aid in complement activation	Most common antibody in the bodily fluids Found in monomeric form Low levels at mucosal sites
IgA	• Specialized for neutralization of targets—weakly induces phagocytosis or complement	The main antibody found at mucosal sites (e.g. intestine and lung) in a dimeric form
IgE	• Specialized for activating Mast cells which can induce rapid responses, such as in allergies	Bound to surface receptors upon mast cells

Antibodies. Proteins, glycoproteins, and polysaccharides on the surfaces of pathogens, are all potential targets for antibodies. The specific area of the macromolecule which is targeted by the antibody is known as the epitope. B cells that have never encountered antigen initially express exclusively IgM and IgD isotypes in a membrane form. Once the cell has become activated through its interaction with a CD4 T cell (described earlier), it will initially release IgM antibodies into the serum in a pentameric form. Over time, the B cell may undergo additional genetic changes in a process known as *class switching* or *isotype switching*. Through switching, the variable region produced by VDJ recombination is matched with a new constant region, altering the function of the antibody. Depending on the specific mix of cytokines, which the B cell receives during activation, it may produce *IgG*, *IgA*, or *IgE* antibody types

be released into the blood where it can specifically bind to and inactivate pathogens (Table 2).

Unlike T cells, activated B cells can further improve the affinity of their immune receptors through further mutation and selection. Those B cells bearing an enhanced receptor will be able to better bind their target pathogen, better recruit associated T cells, and will ultimately receive more activation signals. Thus through successive rounds of target binding, activation, and mutation, a B cell with significantly higher affinity can be selected.

After lymphocytes become activated and the infecting pathogen is neutralized, a majority of the lymphocytes will perish, and about 5% go on to form *immunological memory*. These memory lymphocytes take up residence within lymph node compartments

and can survive there for decades. In the case of reinfection with the same or a similar pathogen, the memory cells react much more quickly, compared to naive lymphocytes, and can yield protective responses within 2–3 days instead of the 7 or more days it takes to generate a primary response.

The cells of the adaptive immune response possess receptors capable of binding to a nearly infinite number of pathogen molecules, but it can also lead to misdirected and harmful immune responses as seen in *autoimmunity*. In some cases T cells can become activated to respond against antigens expressed by host cells. Naturally, an immune response directed against self-antigens is highly undesirable; thus the immune system has developed mechanisms to stop this from happening. One example of this is the negative selection of self-interacting T cells which occurs in the thymus as T cells develop. Despite this and other protective mechanisms, the immune response can still sometimes become misdirected. In the case of rheumatic fever, for example, antibodies developed against a strain of *Streptococcus* bacteria can cross-react with antigens in the heart, which can lead to serious heart damage.

5 Arming the Immune System to Fight Disease

As our understanding of the immune system has grown, so has our ability to use it to generate more effective cancer therapies and vaccine adjuvants. The first vaccines consisted of attenuated or killed pathogens capable of stimulating our immune system effectively to resist future infections but impaired in their ability to cause full-blown infections themselves. More recently subunit vaccines, consisting of specific carbohydrate/protein fractions of the pathogen, are more commonly used as the industry has moved toward better characterized and safer vaccines. However, these subunits are less immunogenic than their whole cell counterparts and require the addition of an adjuvant to stimulate the immune system to mount a sufficiently strong immune response to the vaccine antigen. Aluminum salts (alum) are the original adjuvant, first administered in the early 1930s in toxoid vaccine formulations against *C. tetani* and *C. diphtheria* [22]. Its inclusion in vaccine formulations was shown to induce higher antibody responses to vaccine antigens in vivo. At the time, there was a lack of an in-depth understanding of alum's mechanism of action, but now it is widely thought to have some immunomodulatory activity. More recently, guided by the immune activation that occurs when pattern-recognition receptors are bound by PAMPs, scientists designed synthetic PAMP analogs, such as monophosphoryl lipid A (MPLA) and CpG, to act as more potent vaccine adjuvants [23, 24]. Adjuvants (*see* Table 3) have allowed the development of novel vaccines that protect

Table 3
Adjuvants currently in use for human vaccines

Adjuvant	Type	Vaccine	Pathogen	Year first approved in US
Alum	Aluminum salts such as aluminum hydroxide and aluminum phosphate	Multiple	Hepatitis A & B Human papilloma virus Meningococcus Pneumococcus Pertussis Tetanus Diphtheria	1930s
AS04	TLR4 agonist MPLA adsorbed onto alum	Cervarix®	Human papilloma virus	2009
MF59	Squalene oil-in-water	Fluad®	Influenza	2015
1018 ISS	TLR9 agonist; B class CpG DNA nucleotide	Heplisav-B®	Hepatitis B	2017
AS01	Saponin QS-21 and TLR4 agonist with liposomes	Shingrix®	Herpes zoster (shingles)	2017

against new pathogens and that are more effective in certain patient populations with weakened immune systems (e.g., diabetics).

During the last decade, cancer therapy has undergone a rapid revolution in care with the rapid growth of immune-oncology. Now viewed as a fourth pillar of cancer treatment (joining chemotherapy, surgery, and radiotherapy), immunotherapies which reactivate anti-cancer immune responses have led to significant improvement in survival rates for many types of cancer. The most broadly deployed example of cancer immunotherapies is the so-called checkpoint inhibitors, which function to block the inhibitory signals that often prevent T cells from attacking tumors. In 2010, the first checkpoint inhibitor antibody which blocks CTLA4 (ipilimumab) was approved for the treatment of late stage melanoma; when it was shown that survival rates could be almost doubled, it was added to standard of care treatment [25]. This was followed in 2014 by FDA approval of an antibody targeting the PD-1 checkpoint receptor for the treatment of advanced melanoma [26]. The use of PD-1 and CTLA4 targeting therapy has now become standard of care for late-stage melanoma and is being successfully rolled out for treatment of other hard-to-treat cancers, including but not limited to renal cell carcinoma, lung cancer, head and neck cancers, bladder cancer, and Hodgkin's lymphoma [27].

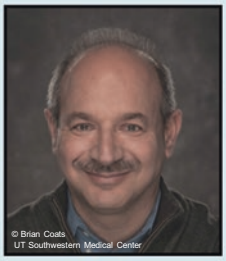
In addition to the exciting introduction of checkpoint inhibitors, a cutting edge genetically modified cell therapy has now cured thousands of patients with previously untreatable forms of B-cell leukemia. Known as chimeric antigen receptor T-cell therapies

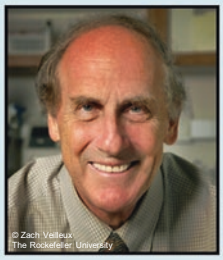
(CAR-T), these therapies work by removing a patient's immune cells and making genetic modifications that allow them to target and kill cancer cells before reinfusing the cells back into the patient [28]. The basis of CAR-T therapy is to create a novel receptor that includes the combination of an antibody that specifically targets cancers (CD19, in the case of B-cell leukemia) and the signaling domains of the T-cell receptors, CD3 and CD28 or 4-1BB. When the antibody portion of the CAR binds to its antigen, this will induce a signal to the T cell to activate both T-cell proliferation and release of cytotoxic proteins from the T cell, leading to killing of the target cell. CAR-T therapies represent the culmination of many threads of basic research over the past decade including HIV and lentiviral research, gene therapy, autologous T-cell therapies, and cancer immunotherapy and were initially considered unlikely to work. The world of cancer therapy was shocked as reports surfaced indicating a complete response rate of over 50% in previously untreatable cases of relapsed and drug-refractory B-cell leukemia, and cure rates are improving as clinicians become increasingly familiar with these therapies [29]. In contrast, CAR-T-cell therapy has not been successful in treating aggressive solid tumors, but researchers around the world are leveraging additional advanced technologies like CRISPR genome editing to try to improve the success of these therapies. In the future it may also be possible to engineer CAR-T-like cells to treat a wide variety of diseases such as autoimmune or degenerative illnesses.

6 Conclusion

Despite the abundance of basic scientific research in the field of immunology, it remains a science intimately bound to practical medical benefit. The development of vaccines is considered to be immunology's greatest success, as evidenced by the millions of lives saved. However, of the estimated 500 known human infectious diseases, we are actively vaccinating against only 26 [30, 31]. While immunologists have had great successes in the past, such as the elimination of smallpox and the near eradication of polio, many more complex diseases continue to challenge us (Fig. 9). To compete with rapidly evolving pathogens like influenza and drug-resistant tuberculosis, we need state of the art technologies to quickly identify the molecular changes in pathogens and immune responses. With modern techniques in proteomics, such as mass spectrometry, it has become possible to quickly identify immune molecules in the lab and the clinic [32]. This technique generates whole proteome results quickly, providing scientists with the data to better understand the mechanisms behind successful immune response so that they can be translated into successful vaccines. It also provides a means to improve upon diagnostics in the clinic by generating a complete







Dr.Jules A. Hoffmann (top), Dr. Bruce A. Beutler (middle) and Dr. Ralph M. Steinman (bottom) split the Nobel Prize in Physiology or Medicine in 2011.

The study of the interactions between the innate and adaptive immune systems has brought about important discoveries in immunology. An important example is the discovery of dendritic cells as the major source of T and B cell activation. With a clearer understanding of how dendritic cells operate, new therapies can be designed to target these cells making it possible to stimulate robust immune responses towards illnesses previously difficult to treat and prevent such as those caused by viruses, intracellular bacteria, and cancer.

In 2011, the Nobel prize for Physiology or Medicine was divided between Dr. Bruce A. Beutler, M.D and Dr. Jules A. Hoffmann, Ph.D., for their discovery of the toll like receptor responsible for innate cell activation and to Dr. Ralph M. Steinman, M.D., for his discovery of the cell which bridges innate and adaptive immunity, the dendritic cell. Dr. Steinman spent his life charting out a path for the use of dendritic cells in combating chronic infections such as tuberculosis, HIV and cancer. When Dr.'s Hoffmann and Beutler discovered that cell surface toll-like receptors were responsible for binding to pathogen debris and activating dendritic cells, driving the cytokine production that guided T and B cell activation, it became possible to create immune modulating designer vaccines. New vaccines can now incorporate toll-like receptor ligands as adjuvants (LPS or ssRNA) that activate dendritic cells in a controlled manner to stimulate T or B cell responses.

These discoveries have also led to a novel approach to cancer therapy, termed dendritic cell immunotherapy. When Dr. Steinman was diagnosed with pancreatic cancer in March of 2007 he teamed up with collaborators around the world to design a dendritic cell based therapy to target his own cancer. Dr Steinman's own dendritic cells were cultured with cytokines *ex vivo*, and exposed to extracts of his tumour. The cytokine groomed dendritic cells, decorated with antigenic peptides from his tumour, were delivered back in the form of a customized vaccine.

Ideally this would have stimulated a robust anti-tumour T cell response. Unfortunately, Dr. Steinman passed away in October of 2011, just 3 days before it was announced he had won the Nobel prize. While it is not known whether the dendritic cell therapy lengthened Dr. Steinman's life, his contributions to science have most certainly advanced the field of cancer immunotherapy forward. In 2010, the first dendritic cell immunotherapy, Sipuleucel-T, was approved to treat prostate cancer. New research into cancer vaccines will continue, drawing on Dr. Steinman's contributions, at the Ralph Steinman center for cancer vaccines at Baylor in Dallas Texas.

Fig. 9 Nobel Prize for Physiology or Medicine winners of 2011

proteomic picture of a person's state of health. The faster we can refine our research methods and unveil protective immune mechanisms in humans, the better chance we stand at developing successful therapies and vaccines for the future.

References

1. Jenner E (1909) The three original publications on vaccination against smallpox. P.F. Collier & Son Company, New York
2. Cavaillon J-M (2011) The historical milestones in the understanding of leukocyte biology initiated by Elie Metchnikoff. J Leukoc Biol 90:413–424
3. Kantha SS (1991) A centennial review; the 1890 tetanus antitoxin paper of von Behring and Kitasato and the related developments. Keio J Med 40:35–39

4. Turk JL (1994) Almroth Wright—phagocytosis and opsonization. J R Soc Med 87:576–577
5. Dreyer WJ, Bennett JC (1965) The molecular basis of antibody formation: a paradox. Proc Natl Acad Sci U S A 54:864–869
6. Murphy K, Travers P, Walport M (2008) The humoral immune response. In: Janeway's immunobiology, 7th edn. Garland Science, New York
7. Feola M, Merklin R, Cho S, Brockman SK (1977) The terminal pathway of the lymphatic system of the human heart. Ann Thorac Surg 24:531–536
8. Medina E (2009) Neutrophil extracellular traps: a strategic tactic to defeat pathogens with potential consequences for the host. J Innate Immun 1:176–180
9. Paust S, von Andrian UH (2011) Natural killer cell memory. Nat Immunol 12:500–508
10. Leavy O (2014) Natural killer cells: RAG keeps natural killers fit. Nat Rev Immunol 14:716–717
11. Karo JM, Sun JC (2015) Novel molecular mechanism for generating NK-cell fitness and memory. Eur J Immunol 45:1906–1915
12. Rapp M, Lau CM, Adams NM et al (2017) Core-binding factor β and Runx transcription factors promote adaptive natural killer cell responses. Sci Immunol 2. https://doi.org/10.1126/sciimmunol.aan3796
13. Berzins SP, Smyth MJ, Baxter AG (2011) Presumed guilty: natural killer T cell defects and human disease. Nat Rev Immunol 11:131–142
14. Bettelli E, Oukka M, Kuchroo VK (2007) T(H)-17 cells in the circle of immunity and autoimmunity. Nat Immunol 8:345
15. Dillman RO (2011) Cancer immunotherapy. Cancer Biother Radiopharm 26:1–64
16. Waldmann TA (2017) Cytokines in cancer immunotherapy. Cold Spring Harb Perspect Biol. https://doi.org/10.1101/cshperspect.a028472
17. Simon JH, Jacobs LD, Campion M et al (1998) Magnetic resonance studies of intramuscular interferon β--1a for relapsing multiple sclerosis. Ann Neurol 43:79–87
18. Asadullah K, Sterry W, Stephanek K et al (1998) IL-10 is a key cytokine in psoriasis. Proof of principle by IL-10 therapy: a new therapeutic approach. J Clin Invest 101:783–794
19. Narula SK, Cutler D, Grint P (1998) Immunomodulation of Crohn's disease by interleukin-10. Agents Actions Suppl 49:57–65
20. Vistnes M, Christensen G, Omland T (2010) Multiple cytokine biomarkers in heart failure. Expert Rev Mol Diagn 10:147–157
21. Lacour S, Gautier J-C, Pallardy M, Roberts R (2005) Cytokines as potential biomarkers of liver toxicity. Cancer Biomark 1:29–39
22. Marrack P, McKee AS, Munks MW (2009) Towards an understanding of the adjuvant action of aluminium. Nat Rev Immunol 9:287–293
23. Garçon N, Di Pasquale A (2017) From discovery to licensure, the Adjuvant System story. Hum Vaccin Immunother 13:19–33
24. Krieg AM (2012) CpG still rocks! Update on an accidental drug. Nucleic Acid Ther 22:77–89
25. Alexander W (2016) The checkpoint immunotherapy revolution: what started as a trickle has become a flood, despite some daunting adverse effects; new drugs, indications, and combinations continue to emerge. P T 41:185
26. Robert C, Ribas A, Wolchok JD et al (2014) Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a phase 1 trial. Lancet 384:1109–1117
27. Decker WK, da Silva RF, Sanabria MH et al (2017) Cancer immunotherapy: historical perspective of a clinical revolution and emerging preclinical animal models. Front Immunol 8:829
28. Maus MV, June CH (2016) Making better chimeric antigen receptors for adoptive T-cell therapy. Clin Cancer Res 22:1875–1884
29. Park JH, Rivière I, Gonen M et al (2018) Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. N Engl J Med 378:449–459
30. (1997) Combined immunization of infants with oral and inactivated poliovirus vaccines: results of a randomized trial in the Gambia, Oman, and Thailand. J Infect Dis 175:S215–S227
31. Osterholm MT, Kelley NS, Manske JM et al (2013) The compelling need for game-changing influenza vaccines an analysis of the influenza vaccine enterprise and recommendations for the future, October 2012. CIDRAP, University of Minnesota, Minneapolis, MN
32. Vidal M, Chan DW, Gerstein M et al (2012) The human proteome - a scientific opportunity for transforming diagnostics, therapeutics, and healthcare. Clin Proteomics 9:6



Chapter 2

Immunoproteomics Methods and Techniques

Kelly M. Fulton, Isabel Baltat, and Susan M. Twine

Abstract

The varied landscape of the adaptive immune response is determined by the peptides presented by immune cells, derived from viral or microbial pathogens or cancerous cells. The study of immune biomarkers or antigens is not new, and classical methods such as agglutination, enzyme-linked immunosorbent assay, or Western blotting have been used for many years to study the immune response to vaccination or disease. However, in many of these traditional techniques, protein or peptide identification has often been the bottleneck. Recent progress in genomics and mass spectrometry have led to many of the rapid advances in proteomics approaches. Immunoproteomics describes a rapidly growing collection of approaches that have the common goal of identifying and measuring antigenic peptides or proteins. This includes gel-based, array-based, mass spectrometry-based, DNA-based, or in silico approaches. Immunoproteomics is yielding an understanding of disease and disease progression, vaccine candidates, and biomarkers. This review gives an overview of immunoproteomics and closely related technologies that are used to define the full set of protein antigens targeted by the immune system during disease.

Key words Immunoproteomics, Mass spectrometry, Antibody, Antigen, Cancer, Infectious disease, SERPA, SEREX, MHC, Epitope

1 Introduction

The landscape of the immune system is constantly changing and is largely determined by the peptides presented on the surface of immune cells. These peptides may be derived from proteins of invading pathogens (whether they be viral, microbial, or parasitic) or host cells (as in cancer or autoimmune disease). Detection and identification of these immune-active proteins or peptides can therefore be investigated using many of the approaches that have been developed for proteomics studies. As an extension of the proteomics field, the term "immunoproteomics" was first used in 2001 [1]. The study of immune biomarkers or antigens is not new; classical methods such as agglutination, enzyme-linked immunosorbent assay (ELISA), and Western blotting have been used for many years to study the immune response to vaccination or disease. Immunoproteomics, however, is yielding an even greater

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_2, © Springer Science+Business Media, LLC, part of Springer Nature 2019

understanding of disease and disease progression. This is also leading to the identification of vaccine candidates and biomarkers.

Mass spectrometry-based proteomics approaches, focussed upon detection, identification, and quantification of proteins and peptides, have become an important component of most immunoproteomics approaches [2]. Classical bottom-up proteomics studies are widely exploited. Here, proteins are typically enriched or fractionated, digested with trypsin or other proteases, and the resulting peptides are analyzed by liquid chromatography coupled to mass spectrometry (LC-MS) [3]. The resulting peptide tandem mass spectra (MS/MS) are then interrogated against translated genome sequences using a variety of search algorithms to give peptide and therefore protein identification.

Rapid innovations in mass spectrometry (MS) for life sciences have given many options for LC-MS analyses, including shotgun MS with data-dependent analysis (DDA), data-independent analyses (DIA), and targeted quantitative analysis such as multiple reaction monitoring (MRM). All have relative advantages and disadvantages. For example, shotgun or DDA analyses rely on algorithms that typically select the top three most intense ions for MS/MS fragmentation. This can introduce problems with reproducibility [4], in particular when targeting very low abundance proteins. Combining peptide fragmentation techniques, including traditional collision-induced dissociation (CID) with electron transfer dissociation (ETD) or higher-energy collisional dissociation (HCD), can help to increase the proportion of peptides identified in a given sample.

Data-independent acquisition (DIA) [5] fragments all ions without preselection and therefore has the potential to provide information about all detected proteins [6]. This requires specialized software for data interpretation and has been used for peptide quantification [5]. MRM and selective reaction monitoring (SRM) allow targeted detection and quantification of a predefined panel of peptides. Targeted approaches exploit triple quadrupole mass spectrometers, whereby a known list of peptide precursor ions is targeted. The ions are selected in the first quadrupole, fragmented in the second, and peptide fragment ions (transitions) monitored in the third quadrupole. When used in combination with a deuterated internal calibrant peptide, this permits the absolute quantification of target peptides [7, 8].

Immunoproteomics continues to benefit from advances in mass spectrometry. Increased sensitivity and resolution of these instruments, combined with more advanced features such as ion mobility, are permitting the identification of lower abundance analytes involved in an immune response. Of particular interest recently is the identification of major histocompatibility complex (MHC) peptides (i.e., T-cell epitopes).

Herein, we focus upon providing a broad overview of immunoproteomics and closely related techniques that are used to study the immune response and their role in disease diagnostics and vaccine development.

2 Immunoproteomics for Characterization of Antibody Targets

One of the two major arms of the adaptive immune system is the humoral immune response. It relies on activated B-cells secreting large amounts of highly specific antibodies, which bind to microbial or cellular targets, either neutralizing them or tagging them for elimination. Antibodies can be generated against microbial invaders and cancer antigens and sometimes misdirected against self-antigens, resulting in autoimmune disease. Many methods have been developed in order to study the antigen targets of the humoral immune response, and in the following section, we provide an overview of the most commonly used. Figure 1 shows a summary overview of these methods.

2.1 Classical Immunoproteomics: Serological Proteome Analysis

One of the most accessible immunoproteomics approaches relies upon two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), which separates proteins based upon orthogonal physical characteristics: isoelectric point (p*I*) and molecular weight. When combined with Western blotting, the technique is commonly known as serological proteome analysis (SERPA). The 2D-PAGE technique was originally developed in the 1970s and with refinements was popularized for use in biochemistry [9]. Early 2D-PAGE studies were hampered by challenges in protein identification and instead used the gel maps to compare protein patterns under different cellular conditions. With the advances in mass spectrometry and genome sequencing in the late 1990s, 2D-PAGE became the mainstay of comparative proteomics studies. 2D-PAGE can be performed in most protein chemistry labs as a matter of routine, and advances in protein staining and image analysis software have made their use accessible to a broad scientific audience. With well-documented disadvantages, including difficulties in resolving very large, small, hydrophobic, or basic proteins, and the dynamic range of protein abundance, 2D-PAGE has been superseded by non-gel-based proteomics approaches. However, 2D-PAGE still offers a number of advantages and remains one of the few techniques that allows high-quality top-down analysis of intact proteins on a proteome-wide scale, including detection of protein posttranslational modifications (PTMs). One of the most overlooked advantages is the ease with which 2D-PAGE can interface with other biochemical techniques. When combined with Western blotting for detection of antigenic proteins and mass spectrometry-based identification of said proteins from in-gel digests, 2D-PAGE provides a powerful approach for antigen identification.

The antigen used in these SERPA studies can be a whole-cell proteome or subproteome (e.g., membrane fraction). 2D-PAGE resolves the majority of proteins in a sample to a single protein spot, giving the potential to readily identify the individual antigenic

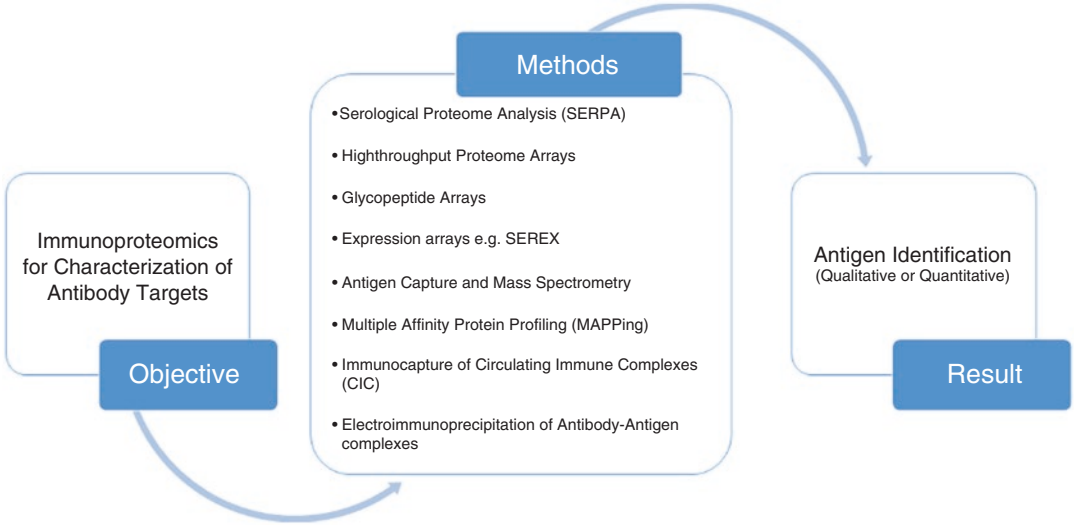


Fig. 1 Overview of methods commonly used to interrogate antigenic targets of the humoral immune response

proteins within the proteome. Gels are then transferred to membranes and probed with sera from animal models or humans and developed as per any traditional [10] Western blotting experiment. Protein-stained gels can be run in parallel to the blotting experiment, providing gels for reference maps and identification of immunoreactive proteins.

In the field of cancer research, SERPA has been used to identify tumor-associated antigens and autoantibodies [11–21] that may serve as biomarkers or lead to the development of therapeutic vaccines. Finally, SERPA is capable of revealing the misdirected immune response that occurs during autoimmune diseases, such as multiple sclerosis [22].

2.2 High Throughput Proteome-Wide Screening of Antibody Targets: The Proteome Array

DNA microarrays were introduced in the mid-1990s, and since then a range of analogous array-based technologies have been developed. These have included arrays for genomic, proteomic, glycomic, and transcriptomic applications. Proteomic applications now include full-length protein, protein fragment, peptide, and antibody arrays. Over the past two decades, protein arrays have become widely used in immunoproteomics, particularly for rapid screening of the humoral immune response [23].

Selection of the protein source is a key decision in the use of protein arrays. Purification of native proteins, from cells or tissues, has the advantage of maintaining native PTMs and potentially protein folding and function. However, purification of single proteins from complex cellular or subcellular mixtures is challenging. The result is often a mixture of unknown proteins, which then requires further deconvolution downstream. Almost all proteome arrays rely on high throughput cloning and recombinant expression of

proteins, with affinity tags for ease of purification. Expression systems include *Escherichia coli*, yeast, insect, or cell-free translation systems. The proteins derived from these expression systems lack native PTMs, and correct protein folding is not guaranteed. In the case of bacterial arrays, the main challenge is the unique repertoire of glycoconjugates and glycoproteins that cannot be replicated by yeast or insect-based expression systems. Recently reported advances in array technology have begun to address more challenging protein antigens, such as membrane proteins [24], and nonprotein antigens, such as carbohydrates [25]. Once purified, proteins are printed onto one of a growing number of fabricated array formats. The arrays are then easily probed with immune serum with binding events detected using fluorescently labeled secondary antibodies [26, 27].

This chip-based technology has several significant advantages; it permits screening of close to equal amounts of antigens, allows interrogation of the entire theoretical proteome of an organism, and requires a comparatively low volume of serum for screening. These benefits, combined with the high throughput capacity of proteome microarrays, make it an attractive method for rapidly screening hundreds of sera. The use of advanced data handling algorithms is a therefore a requirement, as with DNA-based microarrays, for meaningful data interpretation [28, 29].

Infectious disease applications are predominantly focussed upon discovery of diagnostic markers or vaccine antigen candidates. These have included smallpox vaccination [30–32], infections with chlamydia [33, 34], *Mycobacterium tuberculosis* [35–38], *Burkholderia pseudomallei* [39], *Borrelia burgdorferi* [40], *Francisella tularensis* [27, 41], *P. falciparum* [42–45], Epstein-Barr virus [41, 46, 47], Herpes simplex virus infection [48, 49], and *Cytomegalovirus* (CMV) [50], as well as brucellosis [51, 52], salmonellosis [53–55], Q fever [56–58], and toxoplasmosis [59]. For example, the etiological agent of tuberculosis (TB), *M. tuberculosis*, is still poorly understood which hampers the development of vaccines and improvement of diagnostic tests. Proteome microarrays have been fabricated to provide systems level understanding of the pathogen [38]. One study screened serum from 40 patients with latent tuberculosis infection LTBI and active tuberculosis (TB) with a *M. tuberculosis* proteome array and showed a number of protein markers with potential to distinguish between LTBI and TB. Validation of the resulting biomarkers using an ELISA approach moved this one step closer to developing a clinical test [37]. This array based approach also has the potential to be used as a predictor of vaccine efficacy, for example, deciphering immuno-signatures of naturally infected compared with vaccinated individuals. This has been applied, in the study of influenza infection compared to influenza vaccination [60]. More recent applications of the immunoproteome array have included the development

of diagnostic tests to screen the immune response to parasites, such as *Schistosoma mansoni* [61], enterovirus associated with hand, foot, and mouth disease (HFMD) [62], and syphilis [63].

There have recently been intensive efforts to develop human-focussed proteome microarrays, in order to mine the full potential of the human proteome [64]. Much progress has been made in developing human proteome arrays for a variety of purposes, including immunoproteomic profiling for disease biomarker discovery. Examples include ProtoArray [65, 66] and HuProt, which have been used to study the humoral immune response in cancers [67], autoimmune disease [68], and a wide range of infectious diseases. In the long term, this technology has the potential to aid development of improved serodiagnostic tests, vaccine development, and epidemiological studies, as well as shed light on the interaction of pathogens with the host immune system.

2.3 Deciphering the Immune Response to Glycoprotein Antigens

Protein glycosylation is a highly abundant PTM, and aberrant glycosylation of proteins has been shown to be associated with cancers [69] and autoimmune diseases [70, 71]. Truncated glycan moieties on glycoproteins are recognized as nonself and result in the generation of autoantibodies to glycopeptide epitopes [72, 73]. Therefore, immuno-glycoproteomics is essential for vaccine development and the discovery of biomarkers. Glycan microarrays are a technology that has been developed to study native cellular glycan expression, in order to provide insights into how glycan expression plays a role in immunity, signalling, and a myriad of other cellular functions.

Glycans are immobilized or synthesized on a microarray format chip. In the context of immunoproteomics, the resulting chip is screened with immune sera in a conventional high throughput manner. The approach has benefited from improvements in the ability to synthesize a wide repertoire of representative glycans, as well as advancements in chip fabrication and methods of detection [74, 75]. These steps forward have meant that glycan microarrays have become a robust tool in order to gain insight into the response of the immune system toward glycan antigens. In vaccine development, glycan microarrays can be exploited to understand the humoral response to glyco-antigens, the relative antigenicity, markers of successful vaccination, and potential mapping of glycan epitopes [76]. Some recent reports have used glyco-antigen arrays to study the immune response to whole-cell cancer vaccines [76].

O-glycosylation of mucin (MUC1) has been shown to be particularly important in cancers. Patients have been reported to express autoantibodies to distinct epitopes on MUC1 that harbor truncated sugar moieties [77]. As a result of these studies, the MUC1 glycoprotein is considered a target biomarker and a potential vaccine candidate [78]. It is overexpressed by breast, colon, pancreatic, prostate, and ovarian carcinomas. Of note, these auto-

antibodies recognize cancer-specific epitopes, composed of the combined peptide sequence and the carbohydrate moiety [69, 79, 80]. Investigation of glycan-associated autoantibodies has been carried out using variations of chip-based screening technologies, including a microarray display platform that allows the large-scale screening of *O*-glycopeptide libraries for the investigation of disease-associated autoantibodies [73, 81–84]. Rapidly synthesized libraries which represent the potential diversity of glycopeptide or glycoprotein epitopes pave the way to broader screening of glycan epitopes and the elucidation of glycan epitopes within existing immunodominant peptides. There is limited information regarding the clinical levels of autoantibodies in the general population and their potential as biomarkers in screening panels. A 2013 study concluded that autoantibodies to MUC1 glycopeptides could not be used as early detection screening for a range of cancers [85]. Another study used a glycan array and provided evidence that a combination of these signatures have utility in screening for colorectal cancer [86].

The diversity of carbohydrate moieties across the domain of *Bacteria* is substantially greater than that of *Eukarya*. Many monosaccharides are found exclusively within bacteria and are genus, species, or strain specific. Consequently, these unique sugars are often readily identified by the host immune system as foreign entities during infection. Frequently these sugars are part of a pathogen-associated molecular pattern (PAMP), such as lipopolysaccharide (LPS) or peptidoglycan, which is recognized by host pattern recognition receptors (PRR) such as Toll-like receptors (TLR) [87, 88] or nucleotide oligomerization domains (NOD) [89–92] as part of an innate immune response. However, it is increasingly being reported that bacterial glycoproteins also play a role in stimulating innate [93] and adaptive [94–97] host immune responses. Several of these pathogen glycoproteins, including the flagellin of *Campylobacter coli* and *Campylobacter jejuni*, are responsible for serospecific antibody responses [94, 97]. Additionally, the antibody response to anthrose, a unique sugar decorating the *Bacillus anthracis* exosporium glycoproteins (BclA and BclB), is currently being exploited for its potential use in detection and diagnosis of anthrax [98–101]. Despite a growing recognition of the importance of bacterial glycoprotein antigens, immunoproteomics methods directed specifically toward their identification are lacking. In fact, glycoprotein antigenicity is frequently discovered as a consequence of targeted glycoprotein characterization. Given the documented importance of bacterial glycoprotein antigens, methods designed for their global detection and identification would greatly benefit the field of immunoproteomics. Glycan arrays are being exploited as tools for infectious disease research [102], including the development of a carbohydrate vaccine against *Clostridium difficile* [103] and HIV [104]. Synthetic oligomers of the lipoteichoic

acid antigen repeating unit were screened with patient sera to determine immunoreactive epitopes. Glycan microarrays were constructed using components of the capsular polysaccharide of *Streptococcus pneumoniae* 7F [105] and serotype 4 [75] and screened using human sera to determine the minimum protective epitopes for development of a glycoconjugate vaccine [105].

2.4 Antigen Discovery Using Expression Arrays

Expression arrays are composed of bacterial, yeast, mammalian, or cell-free complementary DNA (cDNA) expression libraries that are used to identify novel antigens. These techniques, collectively known as serological analysis of recombinant cDNA expression libraries (SEREX), have a large genetic component and have been termed by some as "reverse proteomics" [106]. SEREX was first developed for analysis of the humoral response to cancer in the 1990s [107], with the goal of identifying tumor-specific antigens that elicit high-titer immunoglobulin G (IgG) antibodies in patient sera. In this context, the technique permits the search for antibody responses and the molecular definition of immunogenic tumor proteins, based upon autologous patient sera (reviewed in 106). Patient tumor mRNA is used to prepare prokaryotically expressed cDNA libraries which are then immunoscreened with absorbed and diluted patient serum for the detection of tumor antigens that have elicited a high-titer IgG humoral response. This approach has the advantage of being able to identify antigens expressed in vivo and is unbiased, based only upon the reactivity of clones with autologous patient serum. A second phase of screening is also carried out, using sera from normal patients in order to define antigens that show cancer-restricted immune recognition [108, 109]. The accuracy of tumor-associated antigen detection is due to the large size of the screening panel [110].

SEREX has been applied to the study of many cancer types, including renal [108, 109, 111, 112], gastrointestinal [113], colon [114–117], and breast [109, 118–130], leading to the identification of cancer-specific antigens. One antigen, NY-ESO-1, was identified in esophageal squamous cell carcinoma, with aberrant expression in various types of malignant tumors [131]. NY-ESO-1 shows restricted expression patterns, elicits both cell-mediated and humoral immune responses [131, 132], and has been under development as a cancer vaccine target. In addition, head and neck squamous cell carcinomas have been studied using a cDNA expression library to identify tumor-associated antigens for immunotherapy [133]. SEREX has also been used to identify potential biomarkers in breast cancer [134], colon cancer [117], and thyroid cancer [135, 136]. Tumor-associated antigens identified from SEREX screening are updated in the Cancer Immunome Database (*see* ref. 137; http://ludwig-sun5.unil.ch/CancerImmunomeDB/). Over 2000 autoantigens are listed in this online database. An excellent review that discusses the classes

of SEREX-defined antigens and the wider impact of this technique upon cancer vaccine and diagnostic development can be found here [106]. SEREX, however, is not limited to the study of cancer. It has also been used to identify antigens associated with stroke and to screen antigens for studies on multiple sclerosis [138] and pancytopenia [139].

Despite many advantages, SEREX presents some challenges, in that it is time consuming to construct cDNA libraries for each tumor sample. In addition, false positives are possible, either due to reactivity with prokaryotic expression components or lack of expression of PTM in prokaryotic expression systems. In particular, protein glycosylation of eukaryotic proteins can be important for antigenic epitopes, including disease-associated changes in glycosylation. A few autoantibodies to PTM-protein epitopes have been reported, including those found in cancers [72, 73] and autoimmune diseases [70]. The use of eukaryotic expression systems can ensure that expressed proteins are glycosylated appropriately [70, 140–143].

2.5 Antigen Capture and Mass Spectrometry

Immunocapture mass spectrometry aims to enrich antigen proteins from complex cell lysates, using mass spectrometry as the final means to identify the captured proteins. There are many variants of immunocapture, and generally immunoglobulins from patient sera are immobilized on protein A or protein G in a column format. This is followed by the application of a cell or tissue lysate to the column, effectively enriching for antigenic proteins (i.e., those proteins to which there are antibodies in patient serum). Captured proteins are then eluted from the column, enzymatically digested, and subsequently identified by MS/MS [144].

2.5.1 Multiple Affinity Protein Profiling

Multiple affinity protein profiling (MAPPing) is an example of an immunocapture technique that has primarily been exploited to identify cancer-related autoantigens [145, 146]. It is based upon two-dimensional immunoaffinity chromatography, whereby antigens from tumor lysates are separated based upon their affinity for immunoglobulins from healthy controls in the first dimension and immunoglobulins from cancer patients in the second. The first dimension removes autoantigens that are recognized by sera from healthy patients. Cancer-restricted autoantigens then flow through to the second column, which then selectively binds them. The proteins eluted from the second chromatography step are therefore likely to be cancer specific and are identified by enzymatic digestion and MS/MS analyses [145, 146].

2.5.2 Capture and Identification of Circulating Immune Complexes

Another variation of immunocapture targets circulating immune complexes (CIC) in serum. Immune complexes (ICs) are formed from the noncovalent interaction between antigens and antibodies and are usually removed by mononuclear phagocytes through

complement receptors and Fc-receptors [147]. This process constantly occurs naturally in healthy individuals and ensures the rapid clearance of denatured proteins, antigens of gut bacteria, or dead cells. Studies have reported that these antigen–antibody complexes can play a role in progression of human autoimmune diseases [147–149], cancer [150], or infectious diseases [151, 152]. Some have argued that identification of antigens incorporated into ICs may be of greater relevance than information regarding free antigens [153]. There is, however, some discrepancy in the literature regarding the utility of IC antigens in disease diagnosis, treatment, or as an indicator of disease severity [153–158]. ICs are isolated from serum [153, 158], using immobilized protein A or G to capture serum antibodies and any associated antigens. The captured antigen–antibody complexes (i.e., ICs) can then be subsequently analyzed using modern proteomic analyses [153, 158–162]. This may involve the direct digestion of the complexes followed by label-free shotgun mass spectrometry. This was applied in the study of advanced glycation end products in stimulating an immune response in hyperglycemic diabetes. The immune complexome analysis showed increased levels of serum albumin, particularly glycated forms [163].

Immune complexome analysis can be applied to many disease states For example, these analyses have been employed to study several major autoimmune diseases, including rheumatoid arthritis (RA) [158], antineutrophil cytoplasmic antibody-associated vasculitis (AAV), Takayasu's arteritis (TA), mixed connective tissue disease (MCTD), dermatomyositis (DM), Sjögren's syndrome (SS), systemic lupus erythematosus (SLE), and systemic scleroderma (SSc), in the search for disease-specific antigens [164]. The study identified 468 distinct IC-associated antigens, with 62 of those found to be disease specific [164]. In the case of rheumatoid arthritis, a recent study identified CICs containing the proteins thrombospondin-1 and platelet factor 4 in the serum of 81% and 52% of a sampling of rheumatoid arthritis patients, respectively [158].

This approach has also been applied to the study of cancer antigens, with the identification of tumor-specific CIC antigens in several cancers, including lung, colon, bladder, renal, and malignant lymphoma [165]. CICs are more readily identifiable compared to antigens embedded in solid tumors and have the potential to reveal antigens that could be used as part of a therapeutic strategy [165]. IC analysis has also been exploited in the study of parasitic diseases. One example is the identification of CICs associated with Chagas disease caused by *Trypanosoma cruzi* [166]. Recently, IC analysis also proved to have utility in the identification of disease markers for diagnosis of infection with *Leishmania donovani* [167]. Immune complexome analysis has also been used to support elucidation of biomarkers of transplant rejection [168].

2.5.3 Electro-immunoprecipitation of Antigen–Antibody Complexes

Electroimmunoprecipitation can exploit differences in electrophoretic mobility between an antibody and its corresponding antigen, resulting in antigen–antibody complexes embedded in an agarose gel. Staining of the gels permits visualization of precipitated complexes. Elution of these complexes, plus enzymatic digestion, and subsequent mass spectrometry analysis can identify the unknown antigenic proteins of interest [169]. In rocket immunoelectrophoresis (RIE), a monoclonal antibody is used. However, crossed immunoelectrophoresis (CIE) involves two dimensions of separation [170, 171] and can therefore be used to identify antigenic proteins reacting with mixtures of monoclonal antibodies, polyclonal antibodies, or serum. Therefore, electroimmunoprecipitation can be used to capture antigens relevant to various disease states or contribute to validation of antigenic proteins. Electroimmunoprecipitation has the added advantage of being quantitative [170–172] and can therefore also be used to monitor the level of serum antibody response to a known antigen.

2.6 Epitope Mapping

Discovery of antigenic proteins is the first step in profiling the humoral immune response to disease. There is often a need to then further dissect the immune response and determine the region of the antigenic protein, i.e., the epitope, which stimulates the immune response. Epitope identification is particularly important to increase understanding of the mechanisms of antibody recognition of antigens, which will support antibody design and epitope-based vaccine design [173]. This can be carried out using a wide variety of techniques, including X-ray crystallography [174], nuclear magnetic resonance (NMR) [175], mass spectrometry [176], hydrogen-deuterium exchange mass spectrometry [177], peptide arrays [178], mutagenesis, and protein cross-linking coupled mass spectrometry [179]. There are also numerous in silico tools available for the prediction of epitopes. A full description of this growing field is beyond the scope of this review, and we direct the reader to recent reviews [180].

3 Immunoproteomics for Characterization of the Cell-Mediated Immune Response

The cell-mediated immunity (CMI) arm of the adaptive immune response involves activation of cell populations such as phagocytes and T-cells and can include the release of communicator molecules, such as cytokines and chemokines in response to foreign invaders or autoantigens. T-cells recognize antigens that are displayed on the surface of host cells in complexes known as the major histocompatibility complex (MHC). The antigens found in complex with MHC molecules are short peptides that are derived from intracellular proteolysis of proteins. This antigen presentation

and processing allows for host recognition of foreign peptides from infected or transformed cells, stimulating an immune response. In addition, this process allows for constant surveillance of peptides derived from the host organism, and self-peptide presentation is involved in T-cell development in the thymus and regulation of self-tolerance. The complement of peptides presented on the cell surface by either MHC class I or II molecules is collectively referred to as the immunopeptidome [181].

3.1 Bioassays for Cytokine Detection and Quantification

Cytokines are low molecular weight secreted proteins, ranging from 6 to 70 kDa [182, 183], with diverse roles in controlling growth, survival, differentiation, and the effector function of cells and tissues (reviewed in refs. [184–186]). They are critical for an effective immune response, and the secreted profiles of certain cells determine the nature of the response (Th1 versus Th2) and dictate whether the immune response is cell mediated or humoral. Production of cytokines is tightly regulated, with an uncontrolled response potentially leading to septic shock. Therefore, controlled production of cytokines is key to many aspects of inflammation and immunity, including a balanced immune response. Therefore, the types and levels of cytokines being expressed can serve as markers of disease progression. There are a growing number of cytokines and closely related growth factors that have been identified [186, 187]. Unlike hormones, cytokines are active over short distances at sites of inflammation and can act in combination with other cytokines to give a variety of biological responses. Cytokine profiles can potentially be indicative of a particular disease state, so in order to correlate this, methods are required that can simultaneously measure levels of multiple cytokines. Although some cytokines are produced at ng/mL concentrations in body fluids, most are expressed at pg/mL levels.

As important biomarkers, there are a range of standard assays that have been used for cytokine detection, including bioassays [188–191], ELISA, enzyme-linked immunospot assays (ELISPOT) [192], and polymerase chain reaction (PCR), which is an indirect measure of cytokine gene expression levels. ELISPOT allows the detection of cytokines from single cells and has been modified to incorporate the sensitivity of a fluorophore (Fluorspot assay). Identifying and quantifying the cytokines secreted in response to a disease state or pathogen are of interest in diagnostics and as vaccine correlates of protection. The cytokine quantification assays that have gained popularity have become increasingly high throughput, allowing for the collection of a greater amount of information about the roles of cytokines during disease or postvaccination. The use of bead-based assays now permits for the multiplex measurement of multiple cytokines simultaneously [193–198].

3.2 Immuno-PCR

Immuno-PCR is a technique that was first reported in 1992 [199] and combines advantages of ELISA-type assays with the sensitivity of PCR. It has utility in the detection of low abundance immune-related protein, such as cytokines [200] and chemokines. As outlined in Fig. 2, the antigen of interest is captured by a specific antibody and in a manner similar to traditional ELISA; a secondary antibody is used to detect binding. In this case, the secondary antibody is a chimeric antibody, with a DNA strand as the detection marker. The incorporation of a DNA tag allows amplification of the detection signal by PCR. This provides many of the advantages of PCR amplification, which are lacking in traditional ELISA assays. Immuno-PCR has been reported to have a 10- to 1000-fold increase in sensitivity compared to traditional antigen detection methods [199, 201], with high potential for the development of diagnostic assays. The technique has reported utility in detection of serological markers of cancer [202, 203], Central nervous system (CNS) indicator proteins [204], detection and quantification of amyloid β-peptide in Alzheimer's disease [204], early diagnosis of infectious diseases [205–210], cytokine detection [211, 212], toxin detection [213–217] and allergy diagnostics [218–220]. In addition, this method is not aimed at discovery of immuno-reactive proteins and has been developed for speed and sensitivity for use as a

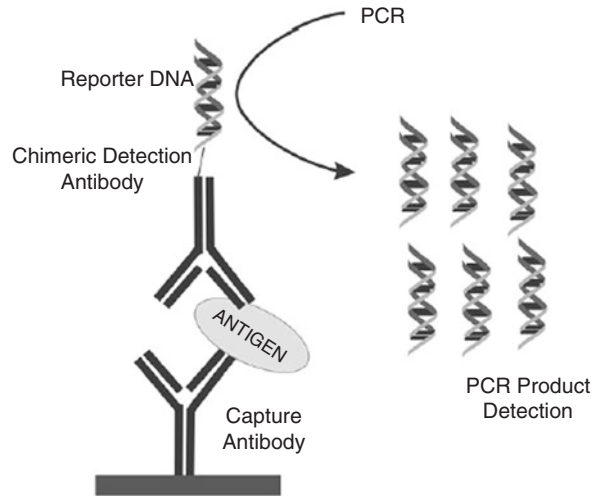


Fig. 2 Immuno-PCR. The setup of immuno-PCR is similar to that of traditional antigen detection ELISA. A capture antibody immobilizes the antigen and detection antibody i added. Instead of the antibody–enzyme conjugate used for colorimetric detection in ELISA, the chimeric antibody with reporter DNA is used. Addition of primers, nucleotides, and polymerase allows amplification of the signal. The linear amplification of PCR means that the number of PCR amplicons generated is proportional to the initial amount of antigen detected. This shows a simplified scheme, and many variations have been developed

clinical laboratory tool [182]. The development of real-time quantitative immuno-PCR has added the ability to measure the amounts of antigen in a sample [183, 221, 222]. An excellent review provides more details on this approach [223–225].

3.3 Immunoaffinity Approaches for Cytokine Detection

These assays are robust, but they are inherently biased toward a predetermined panel of cytokines and provide only quantitative information. Therefore these methods cannot truly be used in discovery research. In addition, these methods provide no information regarding PTM of cytokines, which can be of importance in some cases. For example, IL-24 activity is dependent upon formation of a disulfide bond and glycosylation [226]. Several different immunoproteomics approaches have been reported that are able to detect and quantify cytokines and provide information regarding PTMs. A recently reported technique, known as immunoaffinity capillary electrophoresis (IACE), captures cytokines by immunoaffinity using specific antibodies and then separates the captured proteins using capillary electrophoresis. The resulting protein or peptide fractions are then analyzed by tandem mass spectrometry, providing cytokine identification [227–229]. This two-dimensional separation also allows for differentiation between protein isoforms and identification of PTMs. Another cytokine detection method also exploited an immunoaffinity capture step coupled to analysis by mass spectrometry for protein detection and quantitation [230]. With both methods, the immunoaffinity capture step limits cytokine detection to a predetermined panel. However, in the latter study, the authors' goal was purely to improve the speed of cytokine detection compared to current assay technologies (1–3 h) [230].

3.4 Mass Spectrometry for Cytokine Detection and Quantification

Proteomics and mass spectrometry workflows can be applied to the unbiased detection of cytokines in serum or in vitro secretion from immune cells, such as monocytes. This affords some advantages in discovery studies, compared with using more traditional immunologically based bioassay methods. However, detection of cytokines in serum presents many challenges. Cytokines are typically a very small fraction of the low molecular mass proteome in serum. Although such proteins are amenable to detection using current mass spectrometry technologies, the challenge lies in their low abundance in relation to the high background of other serum proteins. In human serum, albumin and immunoglobulin G (IgG) make up 60–80% of the total serum protein content [231], potentially masking the detection of low abundance proteins. The challenge of the dynamic range of proteins in serum is not new, and there are many strategies for their depletion to improve detection of the more biologically significant low abundance proteins [232]. Additional concerns arise when albumin, known as the "tramp steamer" of the blood, interacts with many small molecules, fatty

acids, and proteins, acting as a transient carrier. Depletion of albumin can then also cause depletion of these transiently bound proteins, peptides, and small molecules which may distort the low abundance serum proteome. Methods have been reported for separation of low molecular weight serum proteins by centrifugal ultrafiltration under solvent conditions that disrupt protein–protein interactions. Two-dimensional liquid chromatography of tryptically digested serum proteins and subsequent identification using mass spectrometry can facilitate the identification of the low MW proteome, including cytokines [232, 233]. Others have used isoelectric focussing (IEF) [234] to improve identification of the low MW serum proteome, while Groessl et al. [235] employed a label-free MS-based proteomics approach to characterize the human monocyte secretome, successfully identifying important proinflammatory proteins and cytokines. Another study compared several serum depletion methods, including nanoparticles, chemical methods, and commercial serum depletion kits [236] as alternatives to traditional immunoaffinity depletion methods [237]. Combining pre-fractionation techniques with MS is essential for cytokine detection. This approach has yielded cytokine identification in cultured cells and secretomes [238]. Several reports have shown DIA analysis to be robust and reproducible in serum biomarker detection [239]. The DIA approach has allowed the identification of serum biomarkers in cancers [240], with one study identifying 568 glycoproteins secreted by a colon adenocarcinoma cell line [241]. Following discovery studies, target MS approaches such as MRM can be exploited for absolute quantification of cytokines in complex serum samples [242]. Advances in these mass spectrometry-based methods are paving the way for rapid, robust, and unbiased serum cytokine detection, characterization, and quantification during disease. This has the potential to contribute to understanding of disease progression, as well as reveal disease or postvaccination biomarkers.

4 Major Histocompatibility Complex Peptides: The Immunopeptidome

There are two major subgroups of MHCs, denoted MHC class I and MHC class II. These two major classes are distinct in their three-dimensional structure, pathways by which antigens are processed, and the type of T-cell with which they interact. The MHC class I gene cluster encodes the heterodimeric proteins that bind antigenic peptide from within cells and are found on all nucleated cell types. MHC class I molecules are carrying peptide antigen complex with the CD8 co-receptor. This complex is primarily recognized by cytotoxic T-cells and leads to their activation and eventual death of the cell expressing the nonself antigen. In comparison, MHC class II gene cluster encodes heterodimeric

peptide-binding proteins and proteins that control peptides binding to the MHC heterodimers. Peptide loading onto MHC class II molecules occurs in the lysosomal pathway, and MHC class II complexes are only found on specialized cell types, such as B-cells, neutrophils, and dendritic cells, and can be induced on macrophages and human T-cells. The CD4 T-cell co-receptor recognizes MHC class II antigen complexes, also resulting in T-cell activation. If the presented peptide is foreign, the T-cells then proliferate, secrete cytokines, and differentiate into antigen-specific effector CD4 cells. These CD4 cells then also secrete cytokines and activate other cell types, such as B-cells which secrete antibodies. Through a series of complex signalling pathways, both the cell-mediated and humoral immune system can be activated by presentation of MHC peptides.

For both MHC class I and class II molecules, the antigens are peptide fragments which are recognized as nonself by T-cells. These antigens are known as T-cell epitopes. A detailed description of how these peptide fragments are generated is described in more detail elsewhere [243–245]. Typically, MHC class I molecules have a binding cleft that accommodates peptides of 8–10 amino acids, whereas MHC class II molecules bind peptides 8–30 amino acids in length. Peptides that bind the cleft of a MHC class II molecule are usually found to share a core sequence [246–251].

In the human population, MHC proteins are encoded by the human leukocyte antigen (HLA) gene clusters. These gene clusters are highly polymorphic, giving rise to hundreds of allelic forms, with only a subset present in each individual. The polymorphism gives rise to differences in the MHC molecules, their binding pockets and affinity for particular peptide antigens, thereby influencing the repertoire of antigens presented to the immune system of an individual. The human genome comprises greater than 10,000 HLA allelic forms, with 6 class I and 8 class II allotypes expressed by each individual, giving rise to enormous complexity of the human immunopeptidome [252].

A deep knowledge of the immunopeptidome is essential for greater understanding of the immune response to disease. The identification and characterization of peptides displayed by MHC molecules and specific T-cell epitopes have become essential for the development of next-generation vaccines and immunotherapeutics against cancer, pathogens, and autoimmunity. A broad array of functional and biochemical approaches have been developed to identify peptide epitopes, including forward and reverse immunoproteomics and mass spectrometry centric approaches.

4.1 MHC Peptide Isolation and Enrichment

Early reports in the 1990s used acid treatment to elute peptides from the surface of cells [253]. Although simple to carry out, peptide elution was not specific to those bound to MHC complexes, and difficulties arose when attempting to discriminate

specific MHC peptides. There are currently three widely adopted approaches: strong acid elution of MHC class I and II peptides from whole-cell lysates using trifluoroacetic acid [254–256], mild acid elution of MHC class I peptides from whole cells [253], and immunoaffinity purification of MHC peptide complexes from cell lysates [257] followed by peptide elution. Typically, relatively large amounts of starting material are required; around 1×10^{10} cells. Figure 3 gives an overview of the current workflow for MHC peptide isolation and identification. Immunoaffinity purification of MHC peptides has been applied in many areas, including the study of the CNS of multiple sclerosis patients [258, 259] and bronchoalveolar lavage cells isolated from patients with sarcoidosis [260]. Another study combines immunoaffinity enrichment with testing of subsequent fractions for biological reactivity, prior to peptide identification by mass spectrometry [261] for the identification of tumor-associated antigens. This approach has also been used for the successful identification of novel antigens in primary human breast cancer [262] and West Nile virus [263].

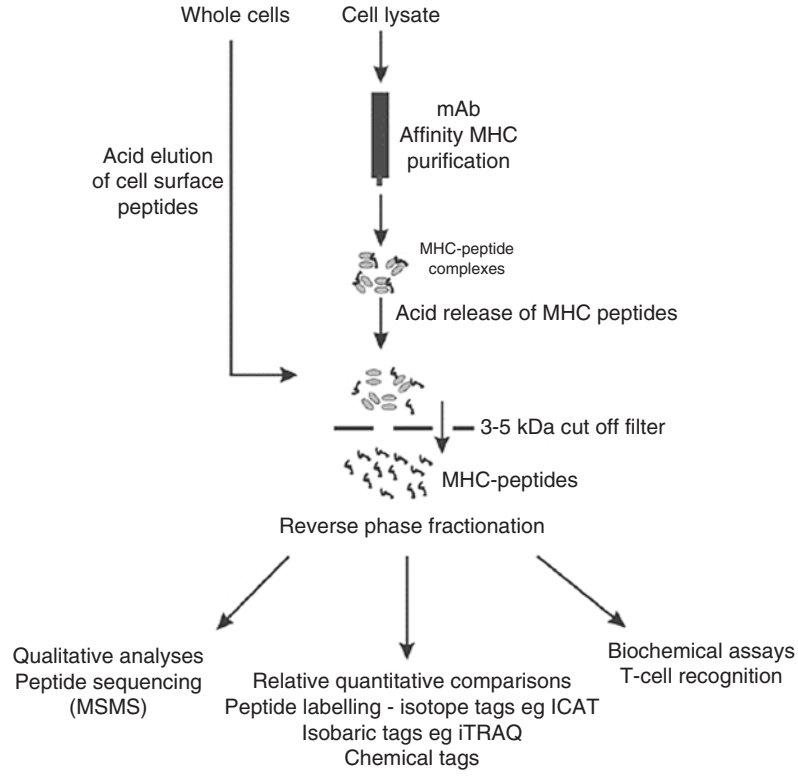


Fig. 3 Schematic overview of methods for MHC peptide purification and sequencing

4.2 Mass Spectrometry Analysis for MHC Peptide Discovery

Pioneering studies in the early 1990s demonstrated the utility of the then recently developed electrospray ionization mass spectrometry (ESI-MS), in combination with microcapillary HPLC to determine the length and sequence of peptides bound to HLA-A2.1 [264], one of the most widely distributed MHC class I molecules within the human population. Since this study, ESI-MS has been used extensively for the detection of peptides presented by MHC molecules in animal models of disease and humans, in a variety of cancers and infectious diseases (e.g., refs. [265–269]). MHC class I and II proteins preferentially bind peptides of different lengths and general characteristics. The challenge associated with MHC peptide identification is distinct from those associated with larger-scale proteome analyses. Immunopeptidome studies typically suffer lower peptide identification rates, due to short peptide lengths and lack of predictable enzymatic cleavage sites [270]. Advances in peptide mass spectrometry over the past decade have increased sensitivity, speed, resolution, and methods of peptide fragmentation, resulting in improved coverage of the immunopeptidome. For example, the use of a combination of mass spectrometry fragmentation techniques, including ETD, HCD, and electron transfer/higher-energy collision dissociation (EThcD), enhances the identification of MHC peptides. One report identified 14,000 MHC class II-bound peptides from matured human monocyte-derived dendritic cells [271].

Reports of the profiling of the MHC II immunopeptidome provide further understanding of the basis for CD4+ T-cell-mediated immunity and tolerance. Studies have reported robust identification of epitopes, and refinements have been made to identify immunodominant epitopes and distinguish self- and nonself-MHC class I peptides. Precise splitting of the eluate from HPLC separation of MHC peptides, with a portion diverted to the mass spectrometer and the majority retained to assay T-cell activity, has allowed more precise correlation between MHC peptide identification and T-cell activation [272–274]. Other methods compared the LC-MS chromatograms of peptides eluted from MHC I complexes with those of reference cells. Mass spectrometry has been used to identify T-cell epitopes of *Plasmodium falciparum* [275], cancers [276, 277], and rheumatoid arthritis [278]. Other studies have combined mass spectrometry and functional assays for T-cell epitope identification [196]. The genetic polymorphism of the HLA alleles results in variation in the MHC complexes across the population, with differing binding affinities. It can, therefore, prove challenging to identify antigenic MHC I peptides presented by MHC class I molecules that are less frequently found across a population. This is important in the development of peptide-based vaccines for the therapeutic treatment of melanoma and other cancers, which requires the identification of antigenic peptides that

will allow the majority of the population, regardless of their MHC-encoded phenotype, to stimulate a T-cell response.

In particular, the field of cancer immunotherapy has experienced a resurgence in recent years, with much interest in personalized cancer vaccines. A number of approaches are exploiting the CD8+ T-cell response against tumor antigens presented on MHC class I complexes. In this context, the combination of improvements in MS technology, informatics approaches, and cancer immunology has been combined so that the field of immunopeptidomics has a high potential impact. There has been a particular focus upon cancer neoantigens, those antigens that are modified in some way from the original antigen, often arising from the tumorigenesis. Neoantigens are unique to tumors and can also be patient specific. These characteristics make them potentially attractive candidates for cancer vaccines [279]. Genomic approaches combined with in silico prediction algorithms have been heavily used to predict point mutations and potential T-cell antigens. The MS-based approaches described have the advantage of providing confirmatory amino acid sequence information, as well as identifying peptides modified with non-template-encoded PTMs. Studies have combined whole-exome sequencing and transcriptome sequencing with MS analysis to identify neoantigens in mouse tumor models [280], with only a small number of the total predicted peptides confirmed by MS to be expressed on MHC I complexes.

Qualitative studies provide an inventory of detected MHC peptides, and with the development of advanced proteomics technologies comes the opportunity to carry out quantitative studies. Quantification of MHC peptides allows for comparison of peptide repertoire and abundance with time, between tissues, individuals, self and nonself, or test and control. Quantification can be relative or absolute and achieved using peptide labeling strategies such as the commercially available isotope coded affinity tag (ICAT) system [281], isobaric tags such as iTRAQ [282], or chemical tags (mass-coded abundance tagging, MCAT) [283]. A recent study, for example, reported robust identification of over 100 MHC II peptides and their relative quantification using stable isotope labeling [284]. SRM was successful in quantifying the amount of a known ovalbumin peptide from the spleens of immunized mice after MHC affinity purification. Recently, the approach has been used to measure the presence and abundance of known MHC melanoma peptide antigens on the surface of several human melanoma cell lines [285]. SRM can be multiplexed for rapid and simultaneous identification and quantitation of hundreds of peptides and is robust and readily transferable between laboratories.

4.3 Characterization of Posttranslational Modifications of T-Cell Epitopes

Glycosylation is a common PTM of proteins in eukaryotes and is increasingly discovered in bacteria. Although largely ignored until recently, carbohydrates, glycolipids, and glycopeptides [286] are now known to modulate T-cell recognition [287, 288] having

been shown to be presented by MHC complexes [289]. This has important implications in the immune response to pathogens, tumor cells, and self-tolerance. Several studies in the late 1990s provided evidence that naturally modified *O*-GlcNAc peptides were ligands for MHC class I molecules [289–292], and a crystal structure showed the glycan moiety to be exposed for recognition by CD8 T-cells [293]. After affinity enrichment of MHC complexes and elution of bound peptides, many of the techniques developed for the study of glycoproteomes could be applied to target and identify MHC bound glycopeptides. Some approaches such as those using lectin enrichment have already been successfully employed for the enrichment of MHC-bound glycopeptides [209]. Other approaches, such as hydrazide capture [294] and chromatographic enrichments, combined with advanced mass spectrometry approaches, such as precursor ion scanning of signature glycan ions, could lead to rapid and specific identification of MHC glycopeptides [295].

Similarly, it has been proposed that phosphopeptides and lipopeptides [296] may also be T-cell antigens [297, 298]. These can be presented by MHC class I molecules on malignant cells [299, 300] or MHC class II [298]. Phosphopeptides associated with class I MHC molecules on the surface of tumor cells can be enriched by immunoaffinity purification of the MHC complexes, followed by elution and enrichment of phosphopeptides with immobilized metal-affinity chromatography (IMAC) [297, 298, 301–303]. If immunogenic, these modified peptides could be attractive targets for cancer immunotherapy [301, 302].

5 Immune Cell Surface Discrimination

MALDI-TOF is seeing increasingly widespread use in clinical microbiology laboratories for the routine identification of bacterial species (e.g., refs. [304–307]). The approach is based upon protein signatures (without protein identification) and exploits not only the differences in cell surface proteins between cell types but the dynamic change in those proteins under certain conditions. This has been demonstrated to be a robust, reproducible, rapid, and potentially cost-saving approach in medical diagnostics [308]. Recently, this approach has been successful in discriminating intact immune cells, including lymphocytes, monocytes, and polymorphonuclear cells, for the generation of an immune cell database [309]. The same approach was also able to distinguish between stimulated and unstimulated macrophages [309]. Further to this, distinct differences in the MALDI-TOF protein fingerprints of the surface of macrophages were detected with the addition of M1 agonists, interferon gamma (IFN-γ), tumor necrosis factor (TNF), lipopolysaccharide (LPS), and LPS + IFN-γ, and the M2 agonists,

interleukin 4 (IL-4), transforming growth factor beta (TGF-β1), and interleukin 10 (IL-10). The differences in macrophage surface fingerprints were specific [310]. A recent study showed this fingerprinting approach to be able to distinguish exogenous activation states of monocyte subpopulations, specifically between monocytes exposed to LPS and *M. tuberculosis* antigens [311].

6 Applications and Future Perspective

Immunoproteomics is becoming a mature field, still with potential impact in the development of immunoproteomics-based assays and the development of vaccines. The breadth and sophistication of the techniques developed to study the immunoproteome have increased dramatically in the past decade. The field has benefited greatly from advances in proteomics and immunoinformatics and will continue to develop. Challenges remain, such as characterization of low abundance T-cell epitopes for the development of cancer vaccines from neoantigens and detection of low-level serum cytokines. However, new avenues of investigation are emerging, including application of interactomics to immunoproteomics studies and comprehensive systems biology studies of the immune response to disease. As our depth of knowledge of the immune response to infection, cancer, or self-antigens (misdirected autoimmunity) increases, so do the opportunities for discovery of robust disease biomarkers for early diagnosis. Combined in silico and experimental studies promise to yield efficacious vaccine candidates and correlates of vaccine protection. On a systems level, understanding the rapidly changing protein landscape of the immune system at various stages of life has the potential to provide immune markers of vaccine health and predictive markers of the immune response, which may in the longer term contribute to the development of personalized medicine.

References

1. Jungblut PR (2001) Proteome analysis of bacterial pathogens. Microbes Infect 3:831–840
2. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198–207
3. Bantscheff M, Lemeer S, Savitski MM, Kuster B (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal Bioanal Chem 404:939–965
4. Tabb DL, Vega-Montoto L, Rudnick PA et al (2009) Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. J Proteome Res 9:761–776
5. Venable JD, Dong M-Q, Wohlschlegel J et al (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. Nat Methods 1:39–45
6. Gillet LC, Navarro P, Tate S et al (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 11:O111.016717
7. Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for

quantitative proteomics: a tutorial. Mol Syst Biol 4:222

8. Picotti P, Rinner O, Stallmach R et al (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods 7:43–46
9. Westermeier R (2016) Electrophoresis in practice: a guide to methods and applications of DNA and protein separations. Wiley, Hoboken, NJ
10. Pajuaba ACAM, Silva DAO, Almeida KC et al (2012) Immunoproteomics of *Brucella abortus* reveals differential antibody profiles between S19-vaccinated and naturally infected cattle. Proteomics 12:820–831
11. Cha HJ, Yoon HG, Kim YW et al (1998) Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. Eur J Biochem 253:251–262
12. Hamrita B, Chahed K, Kabbage M et al (2008) Identification of tumor antigens that elicit a humoral immune response in breast cancer patients' sera by serological proteome analysis (SERPA). Clin Chim Acta 393:95–102
13. Li L, Chen S-H, Yu C-H et al (2008) Identification of hepatocellular-carcinoma-associated antigens and autoantibodies by serological proteome analysis combined with protein microarray. J Proteome Res 7:611–620
14. Forgber M, Gellrich S, Sharav T et al (2009) Proteome-based analysis of serologically defined tumor-associated antigens in cutaneous lymphoma. PLoS One 4:e8376
15. Forgber M, Trefzer U, Sterry W, Walden P (2009) Proteome serological determination of tumor-associated antigens in melanoma. PLoS One 4:e5199
16. Mou Z, He Y, Wu Y (2009) Immunoproteomics to identify tumor-associated antigens eliciting humoral response. Cancer Lett 278:123–129
17. Shukla S, Pranay A, D'Cruz AK et al (2009) Immunoproteomics reveals that cancer of the tongue and the gingivobuccal complex exhibit differential autoantibody response. Cancer Biomark 5:127–135
18. Liu R, Wang K, Yuan K et al (2010) Integrative oncoproteomics strategies for anticancer drug discovery. Expert Rev Proteomics 7:411–429
19. Suzuki A, Iizuka A, Komiyama M et al (2010) Identification of melanoma antigens using a Serological Proteome Approach (SERPA). Cancer Genomics Proteomics 7:17–23
20. Mojtahedi Z, Safaei A, Yousefi Z, Ghaderi A (2011) Immunoproteomics of HER2-positive and HER2-negative breast cancer patients with positive lymph nodes. OMICS 15:409–418
21. O'Meara MM, Disis ML (2011) Therapeutic cancer vaccines and translating vaccinomics science to the global health clinic: emerging applications toward proof of concept. OMICS 15:579–588
22. Almeras L, Lefranc D, Drobecq H et al (2004) New antigenic candidates in multiple sclerosis: identification by serological proteome analysis. Proteomics 4:2184–2194
23. Gupta S, Manubhai KP, Mukherjee S, Srivastava S (2017) Serum profiling for identification of autoantibody signatures in diseases using protein microarrays. Methods Mol Biol 1619:303–315
24. Fang Y, Frutos AG, Lahiri J (2002) Membrane protein microarrays. J Am Chem Soc 124:2394–2395
25. Shin I, Cho JW, Boo DW (2004) Carbohydrate arrays for functional studies of carbohydrates. Comb Chem High Throughput Screen 7:565–574
26. Davies DH, Liang X, Hernandez JE et al (2005) Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc Natl Acad Sci U S A 102:547–552
27. Eyles JE, Unal B, Hartley MG et al (2007) Immunodominant Francisella tularensis antigens identified using proteome microarray. Proteomics 7:2172–2183
28. Sundaresh S, Doolan DL, Hirst S et al (2006) Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques. Bioinformatics 22:1760–1766
29. Díez P, Dasilva N, González-González M et al (2012) Data analysis strategies for protein microarrays. Microarrays (Basel) 1:64–83
30. Davies DH, Molina DM, Wrammert J et al (2007) Proteome-wide analysis of the serological response to vaccinia and smallpox. Proteomics 7:1678–1686
31. Benhnia MR-E-I, Maybeno M, Blum D et al (2013) Unusual features of vaccinia virus extracellular virion form neutralization resistance revealed in human antibody responses to the smallpox vaccine. J Virol 87:1569–1585
32. Pugh C, Keasey S, Korman L, Pittman PR (2014) Human antibody responses to the

polyclonal Dryvax vaccine for smallpox prevention can be distinguished from responses to the monoclonal replacement vaccine ACAM2000. Clin Vaccine Immunol 21(6):877–885

33. Cruz-Fisher MI, Cheng C, Sun G et al (2011) Identification of immunodominant antigens by probing a whole Chlamydia trachomatis open reading frame proteome microarray using sera from immunized mice. Infect Immun 79:246–257
34. Teng A, Cruz-Fisher MI, Cheng C et al (2012) Proteomic identification of immunodominant chlamydial antigens in a mouse model. J Proteome 77:176–186
35. Kunnath-Velayudhan S, Salamon H, Wang H-Y et al (2010) Dynamic antibody responses to the *Mycobacterium tuberculosis* proteome. Proc Natl Acad Sci U S A 107:14703–14708
36. Kunnath-Velayudhan S, Davidow AL, Wang H-Y et al (2012) Proteome-scale antibody responses and outcome of *Mycobacterium tuberculosis* infection in nonhuman primates and in tuberculosis patients. J Infect Dis 206:697–705
37. Cao SH, Chen YQ, Sun Y et al (2018) Screening of serum biomarkers for distinguishing between latent and active tuberculosis using proteome microarray. Biomed Environ Sci 31:515–526
38. Deng J, Bi L, Zhou L et al (2014) *Mycobacterium tuberculosis* proteome microarray for global studies of protein function and immunogenicity. Cell Rep 9:2317–2329
39. Suwannasaen D, Mahawantung J, Chaowagul W et al (2011) Human immune responses to Burkholderia pseudomallei characterized by protein microarray analysis. J Infect Dis 203:1002–1011
40. Barbour AG, Jasinskas A, Kayala MA et al (2008) A genome-wide proteome array reveals a limited set of immunogens in natural infections of humans and white-footed mice with Borrelia burgdorferi. Infect Immun 76:3374–3389
41. Sundaresh S, Randall A, Unal B et al (2007) From protein microarrays to diagnostic antigen discovery: a study of the pathogen Francisella tularensis. Bioinformatics 23:i508–i518
42. Barry AE, Trieu A, Fowkes FJI et al (2011) The stability and complexity of antibody responses to the major surface antigen of *Plasmodium falciparum* are associated with age in a malaria endemic area. Mol Cell Proteomics 10:M111.008326
43. Nnedu ON, O'Leary MP, Mutua D et al (2011) Humoral immune responses to Plasmodium falciparum among HIV-1-infected Kenyan adults. Proteomics Clin Appl 5:613–623
44. Fan Y-T, Wang Y, Ju C et al (2013) Systematic analysis of natural antibody responses to *P. falciparum* merozoite antigens by protein arrays. J Proteome 78:148–158
45. Doolan DL, Mu Y, Unal B et al (2008) Profiling humoral immune responses to *P. falciparum* infection with protein microarrays. Proteomics 8:4680–4694
46. Zheng D, Wan J, Cho YG et al (2011) Comparison of humoral immune responses to Epstein-Barr virus and Kaposi's sarcoma–associated herpesvirus using a viral proteome microarray. J Infect Dis 204:1683–1691
47. Liu Z, Coghill AE, Pfeiffer RM et al (2018) Patterns of interindividual variability in the antibody repertoire targeting proteins across the Epstein-Barr virus proteome. J Infect Dis 217:1923–1931
48. Dasgupta G, Chentoufi AA, Kalantari M et al (2012) Immunodominant "asymptomatic" herpes simplex virus 1 and 2 protein antigens identified by probing whole-ORFome microarrays with serum antibodies from seropositive asymptomatic versus symptomatic individuals. J Virol 86:4358–4369
49. Kalantari-Dehaghi M, Chun S, Chentoufi AA et al (2012) Discovery of potential diagnostic and vaccine antigens in herpes simplex virus 1 and 2 by proteome-wide antibody profiling. J Virol 86:4328–4339
50. Pérez-Bercoff L, Valentini D, Gaseitsiwe S et al (2014) Whole CMV proteome pattern recognition analysis after HSCT identifies unique epitope targets associated with the CMV status. PLoS One 9:e89648
51. Conejero L, Patel N, de Reynal M et al (2011) Low-dose exposure of C57BL/6 mice to Burkholderia pseudomallei mimics chronic human melioidosis. Am J Pathol 179: 270–280
52. Cannella AP, Tsolis RM, Liang L et al (2012) Antigen-specific acquired immunity in human brucellosis: implications for diagnosis, prognosis, and vaccine development. Front Cell Infect Microbiol 2:1
53. Lee S-J, Liang L, Juarez S et al (2012) Identification of a common immune signature in murine and human systemic Salmonellosis. Proc Natl Acad Sci U S A 109:4998–5003
54. Darton TC, Baker S, Randall A et al (2017) Identification of novel serodiagnostic signa-

tures of typhoid fever using a Salmonella proteome array. Front Microbiol 8:1794

55. Liang L, Juarez S, Nga TVT et al (2013) Immune profiling with a Salmonella Typhi antigen microarray identifies new diagnostic biomarkers of human typhoid. Sci Rep 3:1043
56. Beare PA, Chen C, Bouman T et al (2008) Candidate antigens for Q fever serodiagnosis revealed by immunoscreening of a Coxiella burnetii protein microarray. Clin Vaccine Immunol 15:1771–1779
57. Vigil A, Chen C, Jain A et al (2011) Profiling the humoral immune response of acute and chronic Q fever by protein microarray. Mol Cell Proteomics 10:M110.006304
58. Xiong X, Wang X, Wen B et al (2012) Potential serodiagnostic markers for Q fever identified in Coxiella burnetii by immunoproteomic and protein microarray approaches. BMC Microbiol 12:35
59. Liang L, Döşkaya M, Juarez S et al (2011) Identification of potential serodiagnostic and subunit vaccine antigens by antibody profiling of toxoplasmosis cases in Turkey. Mol Cell Proteomics 10:M110.006916
60. Legutki JB, Johnston SA (2013) Immunosignatures can predict vaccine efficacy. Proc Natl Acad Sci U S A 110:18614–18619
61. de Assis RR, Ludolf F, Nakajima R et al (2016) A next-generation proteome array for *Schistosoma mansoni*. Int J Parasitol 46:411–415
62. Zhang A, Xiu B, Zhang H, Li N (2016) Protein microarray-mediated detection of antienterovirus antibodies in serum. J Int Med Res 44:287–296
63. Runina AV, Katunin GL, Filippova MA et al (2018) Immunochip for syphilis serodiagnostics with the use of extended array of Treponema pallidum recombinant antigens. Bull Exp Biol Med 165:767–771
64. Duarte JG, Blackburn JM (2017) Advances in the development of human protein microarrays. Expert Rev Proteomics 14:627–641
65. Schweitzer B, Meng L, Mattoon D, Rai AJ (2010) Immune response biomarker profiling application on ProtoArray® protein microarrays. In: Rai AJ (ed) The urinary proteome: methods and protocols. Humana, Totowa, NJ, pp 243–252
66. Creaney J, Dick IM, Musk AWB et al (2016) Immune response profiling of malignant pleural mesothelioma for diagnostic and prognostic biomarkers. Biomarkers 21:551–561
67. Yang L, Wang J, Li J et al (2016) Identification of serum biomarkers for gastric cancer diagnosis using a human proteome microarray. Mol Cell Proteomics 15:614–623
68. Ayoglu B, Schwenk JM, Nilsson P (2016) Antigen arrays for profiling autoantibody repertoires. Bioanalysis 8:1105–1126
69. Hakomori S (2002) Glycosylation defining cancer malignancy: new wine in an old bottle. Proc Natl Acad Sci U S A 99:10231–10233
70. Coenen D, Verschueren P, Westhovens R, Bossuyt X (2007) Technical and diagnostic performance of 6 assays for the measurement of citrullinated protein/peptide antibodies in the diagnosis of rheumatoid arthritis. Clin Chem 53:498–504
71. Hiki Y (2009) O-linked oligosaccharides of the IgA1 hinge region: roles of its aberrant structure in the occurrence and/or progression of IgA nephropathy. Clin Exp Nephrol 13:415–423
72. Schietinger A, Philip M, Yoshida BA et al (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314:304–308
73. Wandall HH, Blixt O, Tarp MA et al (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res 70:1306–1313
74. Rillahan CD, Paulson JC (2011) Glycan microarrays for decoding the glycome. Annu Rev Biochem 80:797–823
75. Geissner A, Pereira CL, Leddermann M et al (2016) Deciphering antigenic determinants of *Streptococcus pneumoniae* serotype 4 capsular polysaccharide using synthetic oligosaccharides. ACS Chem Biol 11:335–344
76. Padler-Karavani V (2016) Glycan microarray reveal the sweet side of cancer vaccines. Cell Chem Biol 23:1446–1447
77. Smorodin EP, Kurtenkov OA, Shevchuk IN, Tanner RH (2005) The isolation and characterization of human natural alphaGal-specific IgG antibodies applicable to the detection of alphaGal-glycosphingolipids. J Immunoassay Immunochem 26:145–156
78. Cheever MA, Allison JP, Ferris AS et al (2009) The prioritization of cancer antigens: a National Cancer Institute pilot project for the acceleration of translational research. Clin Cancer Res 15:5323–5337
79. Reis CA, David L, Seixas M et al (1998) Expression of fully and under-glycosylated forms of MUC1 mucin in gastric carcinoma. Int J Cancer 79:402–410

80. Brockhausen I (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta 1473:67–95
81. Goldstein DM, Gray NS, Zarrinkar PP (2008) High-throughput kinase profiling as a platform for drug discovery. Nat Rev Drug Discov 7:391–397
82. Blixt O, Cló E, Nudelman AS et al (2011) A high-throughput O-glycopeptide discovery platform for seromic profiling. J Proteome Res 10:1436
83. Kracun SK, Cló E, Clausen H et al (2010) Random glycopeptide bead libraries for seromic biomarker discovery. J Proteome Res 9:6705–6714
84. Pedersen JW, Blixt O, Bennett EP et al (2011) Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. Int J Cancer 128:1860–1871
85. Burford B, Gentry-Maharaj A, Graham R et al (2013) Autoantibodies to MUC1 glycopeptides cannot be used as a screening assay for early detection of breast, ovarian, lung or pancreatic cancer. Br J Cancer 108:2045–2055
86. Pedersen JW, Gentry-Maharaj A, Nøstdal A et al (2014) Cancer-associated autoantibodies to MUC1 and MUC4—a blinded case–control study of colorectal cancer in UK collaborative trial of ovarian cancer screening. Int J Cancer 134:2180–2188
87. Chow JC, Young DW, Golenbock DT et al (1999) Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. J Biol Chem 274:10689–10692
88. Poltorak A, Ricciardi-Castagnoli P, Citterio S, Beutler B (2000) Physical contact between lipopolysaccharide and Toll-like receptor 4 revealed by genetic complementation. Proc Natl Acad Sci U S A 97:2163–2167
89. Girardin SE, Boneca IG, Viala J et al (2003) Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. J Biol Chem 278:8869–8872
90. Girardin SE, Hugot JP, Sansonetti PJ (2003) Lessons from Nod2 studies: towards a link between Crohn's disease and bacterial sensing. Trends Immunol 24:652–658
91. Girardin SE, Travassos LH, Hervé M et al (2003) Peptidoglycan molecular requirements allowing detection by Nod1 and Nod2. J Biol Chem 278:41702–41708
92. Girardin SE, Boneca IG, Carneiro LAM et al (2003) Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. Science 300:1584–1587
93. Verma A, Arora SK, Kuravi SK, Ramphal R (2005) Roles of specific amino acids in the N terminus of *Pseudomonas aeruginosa* flagellin and of flagellin glycosylation in the innate immune response. Infect Immun 73:8237–8246
94. Doig P, Kinsella N, Guerry P, Trust TJ (1996) Characterization of a post-translational modification of Campylobacter flagellin: identification of a sero-specific glycosyl moiety. Mol Microbiol 19:379–387
95. Horn C, Namane A, Pescher P et al (1999) Decreased capacity of recombinant 45/47-kDa molecules (Apa) of *Mycobacterium tuberculosis* to stimulate T lymphocyte responses related to changes in their mannosylation pattern. J Biol Chem 274:32023–32030
96. Romain F, Horn C, Pescher P et al (1999) Deglycosylation of the 45/47-kilodalton antigen complex of *Mycobacterium tuberculosis* decreases its capacity to elicit in vivo or in vitro cellular immune responses. Infect Immun 67:5567–5572
97. Logan SM, Kelly JF, Thibault P et al (2002) Structural heterogeneity of carbohydrate modifications affects serospecificity of Campylobacter flagellins. Mol Microbiol 46:587–597
98. Mehta AS, Saile E, Zhong W et al (2006) Synthesis and antigenic analysis of the BclA glycoprotein oligosaccharide from the *Bacillus anthracis* exosporium. Chemistry 12:9136–9149
99. Wang D, Carroll GT, Turro NJ et al (2007) Photogenerated glycan arrays identify immunogenic sugar moieties of *Bacillus anthracis* exosporium. Proteomics 7:180–184
100. Dhénin SGY, Moreau V, Morel N et al (2008) Synthesis of an anthrose derivative and production of polyclonal antibodies for the detection of anthrax spores. Carbohydr Res 343:2101–2110
101. Dhénin SGY, Moreau V, Nevers M-C et al (2009) Sensitive and specific enzyme immunoassays for antigenic trisaccharide from *Bacillus anthracis* spores. Org Biomol Chem 7:5184–5199
102. Geissner A, Anish C, Seeberger PH (2014) Glycan arrays as tools for infectious disease research. Curr Opin Chem Biol 18:38–45
103. Martin CE, Broecker F, Eller S et al (2013) Glycan arrays containing synthetic *Clostridium difficile* lipoteichoic acid oligomers as tools toward a carbohydrate vaccine. Chem Commun 49:7159–7161
104. Shivatare SS, Shivatare VS, Wu C-Y, Wong C-H (2018) Chemo-enzymatic synthesis of

N-glycans for array development and HIV antibody profiling. J Vis Exp. https://doi.org/10.3791/55855

105. Ménová P, Sella M, Sellrie K et al (2018) Identification of the minimal glycotope of *Streptococcus pneumoniae* 7F capsular polysaccharide using synthetic oligosaccharides. Chemistry 24:4181–4187
106. Lee S-Y, Jeoung D (2007) The reverse proteomics for identification of tumor antigens. J Microbiol Biotechnol 17:879–890
107. Sahin U, Türeci O, Pfreundschuh M (1997) Serological identification of human tumor antigens. Curr Opin Immunol 9:709–716
108. Scanlan MJ, Gordan JD, Williamson B et al (1999) Antigens recognized by autologous antibody in patients with renal-cell carcinoma. Int J Cancer 83:456–464
109. Scanlan MJ, Gout I, Gordon CM et al (2001) Humoral immunity to human breast cancer: antigen definition and quantitative analysis of mRNA expression. Cancer Immun 1:4
110. Vitale M (2013) SEREX: a promising approach for identification of thyroid cancer serological biomarkers. Clin Endocrinol 79:12–13
111. Koroleva EP, Lagarkova MA, Mesheryakov AA et al (2002) Serological identification of antigens associated with renal cell carcinoma. Russ J Immunol 7:229–238
112. Devitt G, Meyer C, Wiedemann N et al (2006) Serological analysis of human renal cell carcinoma. Int J Cancer 118:2210–2219
113. Kobayashi S, Hiwasa T, Arasawa T et al (2018) Identification of specific and common diagnostic antibody markers for gastrointestinal cancers by SEREX screening using testis cDNA phage library. Oncotarget 9:18559–18569
114. Scanlan MJ, Chen YT, Williamson B et al (1998) Characterization of human colon cancer antigens recognized by autologous antibodies. Int J Cancer 76:652–658
115. Line A, Slucka Z, Stengrevics A et al (2002) Characterisation of tumour-associated antigens in colon cancer. Cancer Immunol Immunother 51:574–582
116. Ishikawa T, Fujita T, Suzuki Y et al (2003) Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. Cancer Res 63:5564–5572
117. Garifulin OM, Kykot VO, Gridina NY et al (2015) Application of serex-analysis for identification of human colon cancer antigens. Exp Oncol 37:173–180
118. Obata Y, Takahashi T, Tamaki H et al (1999) Identification of cancer antigens in breast cancer by the SEREX expression cloning method. Breast Cancer 6:305–311
119. Forti S, Scanlan MJ, Invernizzi A et al (2002) Identification of breast cancer-restricted antigens by antibody screening of SKBR3 cDNA library using a preselected patient's serum. Breast Cancer Res Treat 73:245–256
120. Jäger D, Unkelbach M, Frei C et al (2002) Identification of tumor-restricted antigens NY-BR-1, SCP-1, and a new cancer/testis-like antigen NW-BR-3 by serological screening of a testicular library with breast cancer serum. Cancer Immun 2:5
121. Minenkova O, Pucci A, Pavoni E et al (2003) Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. Int J Cancer 106:534–544
122. Jäger D, Taverna C, Zippelius A, Knuth A (2004) Identification of tumor antigens as potential target antigens for immunotherapy by serological expression cloning. Cancer Immunol Immunother 53:144–147
123. Qian F, Odunsi K, Blatt LM et al (2005) Tumor associated antigen recognition by autologous serum in patients with breast cancer. Int J Mol Med 15:137–144
124. Shimada H, Nakashima K, Ochiai T et al (2005) Serological identification of tumor antigens of esophageal squamous cell carcinoma. Int J Oncol 26:77–86
125. Jäger D (2007) Potential target antigens for immunotherapy identified by serological expression cloning (SEREX). In: Sioud M (ed) Target discovery and validation reviews and protocols. Springer, New York, pp 319–326
126. Kiyamova R, Kostianets O, Malyuchik S et al (2010) Identification of tumor-associated antigens from medullary breast carcinoma by a modified SEREX approach. Mol Biotechnol 46:105–112
127. Song M-H, Ha J-C, Lee S-M et al (2011) Identification of BCP-20 (FBXO39) as a cancer/testis antigen from colon cancer patients by SEREX. Biochem Biophys Res Commun 408:195–201
128. Kostianets O, Shyian M, Sergiy D et al (2012) Serological analysis of SEREX-defined medullary breast carcinoma-associated antigens. Cancer Investig 30:519–527
129. Song M-H, Choi K-U, Shin D-H et al (2012) Identification of the cancer/testis antigens AKAP3 and CTp11 by SEREX in hepatocellular carcinoma. Oncol Rep 28:1792–1798

130. Kostianets O, Shyyan M, Antoniuk SV et al (2017) Panel of SEREX-defined antigens for breast cancer autoantibodies profile detection. Biomarkers 22:149–156
131. Chen Y-T, Scanlan MJ et al (1997) A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proc Natl Acad Sci U S A 94:1914–1918
132. Jäger E, Jäger D, Knuth A (1999) CTL-defined cancer vaccines: perspectives for active immunotherapeutic interventions in minimal residual disease. Cancer Metastasis Rev 18:143–150
133. Heubeck B, Wendler O, Bumm K et al (2013) Tumor-associated antigenic pattern in squamous cell carcinomas of the head and neck—analysed by SEREX. Eur J Cancer 49:e1–e7
134. Dyachenko L, Havrysh K, Lytovchenko A et al (2016) Autoantibody response to ZRF1 and KRR1 SEREX antigens in patients with breast tumors of different histological types and grades. Dis Markers 2016:5128720
135. Hou Q, Chen K, Shan Z (2015) The construction of cDNA library and the screening of related antigen of ascitic tumor cells of ovarian cancer. Eur J Gynaecol Oncol 36:590–594
136. Izawa S, Okamura T, Matsuzawa K et al (2013) Autoantibody against WD repeat domain 1 is a novel serological biomarker for screening of thyroid neoplasia. Clin Endocrinol 79:35–42
137. Jongeneel V (2001) Towards a cancer immunome database. Cancer Immun 1:3
138. Muto M, Mori M, Hiwasa T et al (2015) Novel serum autoantibodies against talin1 in multiple sclerosis: possible pathogenetic roles of the antibodies. J Neuroimmunol 284:30–36
139. Hao S, Fu R, Wang H, Shao Z (2017) Screening novel autoantigens targeted by serum IgG autoantibodies in immunorelated pancytopenia by SEREX. Int J Hematol 106:622–630
140. Boder ET, Dane Wittrup K (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15:553–557
141. Mischo A, Wadle A, Wätzig K et al (2003) Recombinant antigen expression on yeast surface (RAYS) for the detection of serological immune responses in cancer patients. Cancer Immun 3:5
142. Wadle A, Mischo A, Imig J et al (2005) Serological identification of breast cancer-related antigens from a *Saccharomyces cerevisiae* surface display library. Int J Cancer 117:104–113
143. Kim M-S, Choi HY, Choi YS et al (2007) Optimized serological isolation of lung-cancer-associated antigens from a yeast surface-expressed cDNA library. J Microbiol Biotechnol 17:993–1001
144. Caron M, Choquet-Kastylevsky G, Joubert-Caron R (2007) Cancer immunomics using autoantibody signatures for biomarker discovery. Mol Cell Proteomics 6:1115–1122
145. Hardouin J, Lasserre J-P, Canelle L et al (2007) Usefulness of autoantigens depletion to detect autoantibody signatures by multiple affinity protein profiling. J Sep Sci 30:352–358
146. Hardouin J-P, Lasserre J, Sylvius L et al (2007) Cancer immunomics: from serological proteome analysis to multiple affinity protein profiling. Ann N Y Acad Sci 1107:223–230
147. Roozendaal R, Carroll MC (2007) Complement receptors CD21 and CD35 in humoral immunity. Immunol Rev 219:157–166
148. Solomon S, Kassahn D, Illges H (2005) The role of the complement and the Fc gamma R system in the pathogenesis of arthritis. Arthritis Res Ther 7:129–135
149. Aibara N, Ichinose K, Baba M et al (2018) Proteomic approach to profiling immune complex antigens in cerebrospinal fluid samples from patients with central nervous system autoimmune diseases. Clin Chim Acta 484:26–31
150. Croce MV, Fejes M, Riera N et al (1985) Clinical importance of circulating immune complexes in human acute lymphoblastic leukemia. Cancer Immunol Immunother 20:91–95
151. Liu P, Overman RG, Yates NL et al (2011) Dynamic antibody specificities and virion concentrations in circulating immune complexes in acute to chronic HIV-1 infection. J Virol 85:11196–11207
152. Hadi SA, Waters WR, Palmer M et al (2018) Development of a multidimensional proteomic approach to detect circulating immune complexes in cattle experimentally infected with Mycobacterium bovis. Front Vet Sci 5:141
153. Ohyama K, Kuroda N (2012) Proteomic approaches to profiling the humoral immune response and identifying disease-associated antigens. Biol Pharm Bull 35:1409–1412
154. Coppo R, Bosticardo GM, Basolo B et al (1982) Clinical significance of the detection

of circulating immune complexes in lupus nephritis. Nephron 32:320–328

155. Soltis RD, Hasz DE (1983) The effect of serum immunoglobulin concentration on immune complex detection by polyethylene glycol. J Immunol Methods 57:275–282
156. Valentijn RM, van Overhagen H, Hazevoet HM et al (1985) The value of complement and immune complex determinations in monitoring disease activity in patients with systemic lupus erythematosus. Arthritis Rheum 28:904–913
157. Lock RJ, Unsworth DJ (2000) Measurement of immune complexes is not useful in routine clinical practice. Ann Clin Biochem 37:253–261
158. Ohyama K, Ueki Y, Kawakami A et al (2011) Immune complexome analysis of serum and its application in screening for immune complex antigens in rheumatoid arthritis. Clin Chem 57:905–909
159. Ohyama K, Kawakami A, Tamai M et al (2012) Serum immune complex containing thrombospondin-1: a novel biomarker for early rheumatoid arthritis. Ann Rheum Dis 71:1916–1917
160. Ohyama K, Kuroda N (2013) Immune complexome analysis. Adv Clin Chem 60:129–141
161. Aibara N, Kamohara C, Chauhan AK et al (2018) Selective, sensitive and comprehensive detection of immune complex antigens by immune complexome analysis with papain-digestion and elution. J Immunol Methods 461:85–90
162. Baba M, Ohyama K, Kishikawa N, Kuroda N (2013) Optimization of separation and digestion conditions in immune complexome analysis. Anal Biochem 443:181–186
163. Bhat S, Jagadeeshaprasad MG, Patil YR et al (2016) Proteomic insight reveals elevated levels of albumin in circulating immune complexes in diabetic plasma. Mol Cell Proteomics 15:2011–2020
164. Ohyama K, Baba M, Tamai M et al (2015) Proteomic profiling of antigens in circulating immune complexes associated with each of seven autoimmune diseases. Clin Biochem 48:181–185
165. Ohyama K, Yoshimi H, Aibara N et al (2017) Immune complexome analysis reveals the specific and frequent presence of immune complex antigens in lung cancer patients: a pilot study. Int J Cancer 140:370–380
166. Ohyama K, Huy NT, Yoshimi H et al (2016) Proteomic profile of circulating immune complexes in chronic Chagas disease. Parasite Immunol 38:609–617
167. Jamal F, Shivam P, Kumari S et al (2017) Identification of Leishmania donovani antigen in circulating immune complexes of visceral leishmaniasis subjects for diagnosis. PLoS One 12:e0182474
168. Aibara N, Ohyama K, Hidaka M et al (2018) Immune complexome analysis of antigens in circulating immune complexes from patients with acute cellular rejection after living donor liver transplantation. Transpl Immunol 48:60–64
169. Beyer NH, Schou C, Houen G, Heegaard NHH (2008) Extraction and identification of electroimmunoprecipitated proteins from agarose gels. J Immunol Methods 330:24–33
170. Grubb AO (1974) Crossed immunoelectrophoresis and electroimmunoassay of IgM. J Immunol 112:1420–1425
171. Grubb AO (1974) Crossed immunoelectrophoresis and electroimmunoassay of IgG. J Immunol 113:343–347
172. Laurell CB (1966) Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. Anal Biochem 15:45–52
173. Gershoni JM, Roitburd-Berman A, Siman-Tov DD et al (2007) Epitope mapping. BioDrugs 21:145–156
174. Malito E, Carfi A, Bottomley MJ (2015) Protein crystallography in vaccine research and development. Int J Mol Sci 16:13106–13140
175. Becker W, Bhattiprolu KC, Gubensäk N, Zangger K (2018) Investigating protein-ligand interactions by solution nuclear magnetic resonance spectroscopy. ChemPhysChem 19:895–906
176. Opuni KFM, Al-Majdoub M, Yefremova Y et al (2018) Mass spectrometric epitope mapping. Mass Spectrom Rev 37:229–241
177. Oganesyan I, Lento C, Wilson DJ (2018) Contemporary hydrogen deuterium exchange mass spectrometry. Methods 144:27–42
178. Forsström B, Axnäs BB, Stengele K-P et al (2014) Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays. Mol Cell Proteomics 13:1585–1597
179. Pimenova T, Nazabal A, Roschitzki B et al (2008) Epitope mapping on bovine prion protein using chemical cross-linking and mass spectrometry. J Mass Spectrom 43:185–195
180. Potocnakova L, Bhide M, Pulzova LB (2016) An introduction to B-cell epitope mapping

and in silico epitope prediction. J Immunol Res 2016:6760830

181. Caron E, Kowalewski DJ, Chiek Koh C et al (2015) Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. Mol Cell Proteomics 14:3105–3117
182. Adler M (2005) Immuno-PCR as a clinical laboratory tool. Adv Clin Chem 39: 239–292
183. Adler M, Wacker R, Niemeyer CM (2003) A real-time immuno-PCR assay for routine ultrasensitive quantification of proteins. Biochem Biophys Res Commun 308: 240–250
184. Dinarello CA (2007) Historical insights into cytokines. Eur J Immunol 37:S34–S45
185. Dinarello CA (2010) Anti-inflammatory agents: present and future. Cell 140: 935–950
186. Stenken JA, Poschenrieder AJ (2015) Bioanalytical chemistry of cytokines—a review. Anal Chim Acta 853:95–115
187. Bruserud O (2010) The chemokine system in experimental and clinical hematology. Springer Science & Business Media, Berlin
188. Thorpe R, Wadhwa M, Bird CR, Mire-Sluis AR (1992) Detection and measurement of cytokines. Blood Rev 6:133–148
189. Whiteside TL (2003) Chapter 61 - Assays for cytokines. In: Thomson AW, Lotze MT (eds) The cytokine handbook, 4th edn. Academic, London, pp 1375–1396
190. Mire-Sluis AR, Page L, Thorpe R (1995) Quantitative cell line based bioassays for human cytokines. J Immunol Methods 187:191–199
191. House RV (1999) Cytokine bioassays: an overview. Dev Biol Stand 97:13–19
192. Favre N, Bordmann G, Rudin W (1997) Comparison of cytokine measurements using ELISA, ELISPOT and semi-quantitative RT-PCR. J Immunol Methods 204:57–66
193. Leng SX, McElhaney JE, Walston JD et al (2008) ELISA and multiplex technologies for cytokine measurement in inflammation and aging research. J Gerontol A Biol Sci Med Sci 63:879–884
194. Yu X, Schneiderhan-Marra N, Hsu H-Y et al (2009) Protein microarrays: effective tools for the study of inflammatory diseases. Methods Mol Biol 577:199–214
195. Han K-C, Ahn D-R, Yang EG (2010) An approach to multiplexing an immunosorbent assay with antibody-oligonucleotide conjugates. Bioconjug Chem 21:2190–2196
196. Szarka A, Rigó J Jr, Lázár L et al (2010) Circulating cytokines, chemokines and adhesion molecules in normal pregnancy and preeclampsia determined by multiplex suspension array. BMC Immunol 11:59
197. Yu X, Hartmann M, Wang Q et al (2010) μFBI: a microfluidic bead-based immunoassay for multiplexed detection of proteins from a μL sample volume. PLoS One 5. https://doi.org/10.1371/journal.pone.0013125
198. Han K-C, Yang EG, Ahn D-R (2012) A highly sensitive, multiplex immunoassay using gold nanoparticle-enhanced signal amplification. Chem Commun 48:5895–5897
199. Niemeyer CM, Adler M, Wacker R (2007) Detecting antigens by quantitative immuno-PCR. Nat Protoc 2:1918–1930
200. Adler M, Wacker R, Niemeyer CM (2008) Sensitivity by combination: immuno-PCR and related technologies. Analyst 133:702–718
201. Sano T, Smith CL, Cantor CR (1992) Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. Science 258:120–122
202. Kasai N, Kobayashi K, Shioya S et al (2012) Soluble heparin-binding EGF-like growth factor (HB-EGF) detected by newly developed immuno-PCR method is a clear-cut serological biomarker for ovarian cancer. Am J Transl Res 4:415–421
203. Chia-Ching L, Subramaniam S, Sivasubramanian S, Feng-Huei L (2016) MWCNT-Fe3O4-based immuno-PCR for the early screening of nasopharyngeal carcinoma. Mater Sci Eng C Mater Biol Appl 61:422–428
204. Kuczius T, Becker K, Fischer A, Zhang W (2012) Simultaneous detection of three CNS indicator proteins in complex suspensions using a single immuno-PCR protocol. Anal Biochem 431:4–10
205. Malou N, Renvoise A, Nappez C, Raoult D (2012) Immuno-PCR for the early serological diagnosis of acute infectious diseases: the Q fever paradigm. Eur J Clin Microbiol Infect Dis 31:1951–1960
206. Mehta PK, Dahiya B, Sharma S et al (2017) Immuno-PCR, a new technique for the serodiagnosis of tuberculosis. J Microbiol Methods 139:218–229
207. Xie Q, Zhang J, Shao H et al (2016) Development of a novel immuno-PCR for detection of avian leukosis virus. J Virol Methods 236:25–28
208. Mehta PK, Singh N, Dharra R et al (2016) Diagnosis of tuberculosis based on the detec-

tion of a cocktail of mycobacterial antigen 85B, ESAT-6 and cord factor by immuno-PCR. J Microbiol Methods 127:24–27

209. Halpern MD, Molins CR, Schriefer M, Jewett MW (2014) Simple objective detection of human lyme disease infection using immuno-PCR and a single recombinant hybrid antigen. Clin Vaccine Immunol 21:1094–1105
210. Mehta PK, Raj A, Singh NP, Khuller GK (2014) Detection of potential microbial antigens by immuno-PCR (PCR-amplified immunoassay). J Med Microbiol 63:627–641
211. Potůčková L, Franko F, Bambousková M, Dráber P (2011) Rapid and sensitive detection of cytokines using functionalized gold nanoparticle-based immuno-PCR, comparison with immuno-PCR and ELISA. J Immunol Methods 371:38–47
212. Assumpção ALFV, da Silva RC (2016) Immuno-PCR in cancer and non-cancer related diseases: a review. Vet Q 36:63–70
213. Chao H-Y, Wang Y-C, Tang S-S, Liu H-W (2004) A highly sensitive immuno-polymerase chain reaction assay for Clostridium botulinum neurotoxin type A. Toxicon 43:27–34
214. Allen RC, Rogelj S, Cordova SE, Kieft TL (2006) An immuno-PCR method for detecting *Bacillus thuringiensis* Cry1Ac toxin. J Immunol Methods 308:109–115
215. Fischer A, von Eiff C, Kuczius T et al (2007) A quantitative real-time immuno-PCR approach for detection of staphylococcal enterotoxins. J Mol Med 85:461–469
216. Zhang W, Bielaszewska M, Pulz M et al (2008) New immuno-PCR assay for detection of low concentrations of shiga toxin 2 and its variants. J Clin Microbiol 46:1292–1297
217. He X, Qi W, Quiñones B et al (2011) Sensitive detection of Shiga toxin 2 and some of its variants in environmental samples by a novel immuno-PCR assay. Appl Environ Microbiol 77:3558–3564
218. Simonova MA, Pivovarov VD, Ryazantsev DY et al (2018) Comparative diagnostics of allergy using quantitative immuno-PCR and ELISA. Bioanalysis 10:757–767
219. Rahmatpour S, Khan AH, Nasiri Kalmarzi R et al (2017) Application of immuno-PCR assay for the detection of serum IgE specific to Bermuda allergen. Mol Cell Probes 32:1–4
220. Kolesnikov AV, Kozyr AV, Ryabko AK, Shemyakin IG (2016) Ultrasensitive detection of protease activity of anthrax and botulinum toxins by a new PCR-based assay. Pathog Dis 74:ftv112
221. Lind K, Kubista M (2005) Development and evaluation of three real-time immuno-PCR assemblages for quantification of PSA. J Immunol Methods 304:107–116
222. Niemeyer CM, Adler M, Wacker R (2005) Immuno-PCR: high sensitivity detection of proteins by nucleic acid amplification. Trends Biotechnol 23:208–216
223. Malou N, Raoult D (2011) Immuno-PCR: a promising ultrasensitive diagnostic method to detect antigens and antibodies. Trends Microbiol 19:295–302
224. Ryazantsev DY, Voronina DV, Zavriev SK (2016) Immuno-PCR: achievements and perspectives. Biochemistry 81:1754–1770
225. Chang L, Li J, Wang L (2016) Immuno-PCR: an ultrasensitive immunoassay for biomolecular detection. Anal Chim Acta 910:12–24
226. Fuson KL, Zheng M, Craxton M et al (2009) Structural mapping of post-translational modifications in human interleukin-24: role of N-linked glycosylation and disulfide bonds in secretion and activity. J Biol Chem 284:30526–30533
227. Guzman NA, Phillips TM (2005) Immunoaffinity CE for proteomics studies. Anal Chem 77:60A–67A
228. Guzman NA, Blanc T, Phillips TM (2008) Immunoaffinity capillary electrophoresis as a powerful strategy for the quantification of low-abundance biomarkers, drugs, and metabolites in biological matrices. Electrophoresis 29:3259–3278
229. Guzman NA, Phillips TM (2011) Immunoaffinity capillary electrophoresis: a new versatile tool for determining protein biomarkers in inflammatory processes. Electrophoresis 32:1565–1578
230. Boyle M, Hess J, Nuara A, Buller R (2006) Application of immunoproteomics to rapid cytokine detection. Methods 38: 342–350
231. Hortin GL, Sviridov D, Anderson NL (2008) High-abundance polypeptides of the human plasma proteome comprising the top 4 logs of polypeptide abundance. Clin Chem 54:1608–1616
232. Björhall K, Miliotis T, Davidsson P (2005) Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. Proteomics 5:307–317
233. Tirumalai RS, Chan KC, Prieto DA et al (2003) Characterization of the low molecular

weight human serum proteome. Mol Cell Proteomics 2:1096–1103

234. Harper RG, Workman SR, Schuetzner S et al (2004) Low-molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry. Electrophoresis 25:1299–1306
235. Groessl M, Luksch H, Rösen-Wolff A et al (2012) Profiling of the human monocytic cell secretome by quantitative label-free mass spectrometry identifies stimulus-specific cytokines and proinflammatory proteins. Proteomics 12:2833–2842
236. de Jesus JR, da Silva Fernandes R, de Souza Pessôa G et al (2017) Depleting high-abundant and enriching low-abundant proteins in human serum: an evaluation of sample preparation methods using magnetic nanoparticle, chemical depletion and immunoaffinity techniques. Talanta 170:199–209
237. Wiederin J, Ciborowski P (2016) 6 - Immunoaffinity depletion of highly abundant proteins for proteomic sample preparation. In: Ciborowski P, Silberring J (eds) Proteomic profiling and analytical chemistry, 2nd edn. Elsevier, Boston, pp 101–114
238. Alečković M, Wei Y, LeRoy G et al (2017) Identification of Nidogen 1 as a lung metastasis protein through secretome analysis. Genes Dev 31:1439–1455
239. Collins BC, Hunter CL, Liu Y et al (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. Nat Commun 8:291
240. Luo Y, Mok TS, Lin X et al (2017) SWATH-based proteomics identified carbonic anhydrase 2 as a potential diagnosis biomarker for nasopharyngeal carcinoma. Sci Rep 7:41191
241. Lin Q, Lim HSR, Lin HL et al (2015) Analysis of colorectal cancer glyco-secretome identifies laminin β-1 (LAMB1) as a potential serological biomarker for colorectal cancer. Proteomics 15:3905–3920
242. Percy AJ, Chambers AG, Yang J et al (2014) Advances in multiplexed MRM-based protein biomarker quantitation toward clinical utility. Biochim Biophys Acta 1844:917–926
243. Hugues S, Malherbe L, Filippi C, Glaichenhaus N (2002) Generation and use of alternative multimers of peptide/MHC complexes. J Immunol Methods 268:83–92
244. Watts C, Moss CX, Mazzeo D et al (2003) Creation versus destruction of T cell epitopes in the class II MHC pathway. Ann N Y Acad Sci 987:9–14
245. Saunders PM, van Endert P (2011) Running the gauntlet: from peptide generation to antigen presentation by MHC class I. Tissue Antigens 78:161–170
246. Rudensky AY, Preston-Hurlburt P, Hong SC et al (1991) Sequence analysis of peptides bound to MHC class II molecules. Nature 353:622–627
247. Rudensky AY, Preston-Hurlburt P, al-Ramadi BK et al (1992) Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. Nature 359:429–431
248. Engelhard VH (1994) Structure of peptides associated with MHC class I molecules. Curr Opin Immunol 6:13–23
249. Engelhard VH (1994) Structure of peptides associated with class I and class II MHC molecules. Annu Rev Immunol 12:181–207
250. Alfonso C, Karlsson L (2000) Nonclassical MHC class II molecules. Annu Rev Immunol 18:113–142
251. Santambrogio L, Strominger JL (2006) The ins and outs of MHC class II proteins in dendritic cells. Immunity 25:857–859
252. Cole DK (2015) The ultimate mix and match: making sense of HLA alleles and peptide repertoires. Immunol Cell Biol 93:515–516
253. Storkus WJ, Zeh HJ III, Salter RD, Lotze MT (1993) Identification of T-cell epitopes: rapid isolation of class I-presented peptides from viable cells by mild acid elution. J Immunother Emphasis Tumor Immunol 14:94–103
254. Rotzschke O, Falk K, Wallny HJ, Faath S (1990) Characterization of naturally occurring minor histocompatibility peptides including H-4 and HY. Science 249(4966):283–287
255. Rötzschke O, Falk K, Deres K et al (1990) Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. Nature 348:252–254
256. Falk K, Rötzschke O, Rammensee HG (1990) Cellular peptide composition governed by major histocompatibility complex class I molecules. Nature 348:248–251
257. Van Bleek GM, Nathenson SG (1990) Isolation of an endogenously processed immunodominant viral peptide from the class I H–2Kb molecule. Nature 348:213–216
258. Duyar H, Dengjel J, de Graaf KL et al (2005) Peptide motif for the rat MHC class II molecule RT1.Da: similarities to the multiple sclerosis-associated HLA-DRB1∗1501 molecule. Immunogenetics 57:69–76

259. Fissolo N, Haag S, de Graaf KL et al (2009) Naturally presented peptides on major histocompatibility complex I and II molecules eluted from central nervous system of multiple sclerosis patients. Mol Cell Proteomics 8:2090–2101
260. Wahlström J, Dengjel J, Persson B et al (2007) Identification of HLA-DR–bound peptides presented by human bronchoalveolar lavage cells in sarcoidosis. J Clin Invest 117:3576–3582
261. Demine R, Sherev T, Walden P (2003) Biochemical determination of natural tumor-associated T-cell epitopes. Mol Biotechnol 25:71–78
262. Verma B, Hawkins OE, Neethling FA et al (2010) Direct discovery and validation of a peptide/MHC epitope expressed in primary human breast cancer cells using a TCRm monoclonal antibody with profound antitumor properties. Cancer Immunol Immunother 59:563–573
263. McMurtrey CP, Lelic A, Piazza P et al (2008) Epitope discovery in West Nile virus infection: identification and immune recognition of viral epitopes. Proc Natl Acad Sci U S A 105:2981–2986
264. Hunt DF, Michel H, Dickinson TA et al (1992) Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. Science 256:1817–1820
265. Lemmel C, Weik S, Eberle U et al (2004) Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling. Nat Biotechnol 22:450–454
266. Lanoix J, Durette C, Courcelles M et al (2018) Comparison of the MHC I immunopeptidome repertoire of B-cell lymphoblasts using two isolation methods. Proteomics 18:e1700251
267. Jappe EC, Kringelum J, Trolle T, Nielsen M (2018) Predicted MHC peptide binding promiscuity explains MHC class I "hotspots" of antigen presentation defined by mass spectrometry eluted ligand data. Immunology 154:407–417
268. Rozanov DV, Rozanov ND, Chiotti KE et al (2018) MHC class I loaded ligands from breast cancer cell lines: a potential HLA-I-typed antigen collection. J Proteome 176:13–23
269. Murphy JP, Konda P, Kowalewski DJ et al (2017) MHC-I ligand discovery using targeted database searches of mass spectrometry data: implications for T-cell immunotherapies. J Proteome Res 16:1806–1816
270. Escobar H, Reyes-Vargas E, Jensen PE et al (2011) Utility of characteristic QTOF MS/MS fragmentation for MHC class I peptides. J Proteome Res 10:2494–2507
271. Mommen GPM, Marino F, Meiring HD et al (2016) Sampling from the proteome to the HLA-DR ligandome proceeds via high specificity. Mol Cell Proteomics 15(4):1412–1423
272. Henderson RA, Cox AL, Sakaguchi K et al (1993) Direct identification of an endogenous peptide recognized by multiple HLA-A2.1-specific cytotoxic T cells. Proc Natl Acad Sci U S A 90:10275–10279
273. Cox AL, Skipper J, Chen Y et al (1994) Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. Science 264:716–719
274. den Haan JM, Sherman NE, Blokland E et al (1995) Identification of a graft versus host disease-associated human minor histocompatibility antigen. Science 268:1476–1480
275. Carralot J-P, Lemmel C, Stevanovic S, Pascolo S (2008) Mass spectrometric identification of an HLA-A*0201 epitope from Plasmodium falciparum MSP-1. Int Immunol 20:1451–1456
276. Hawkins OE, Vangundy RS, Eckerd AM et al (2008) Identification of breast cancer peptide epitopes presented by HLA-A*0201. J Proteome Res 7:1445–1457
277. Weinzierl AO, Maurer D, Altenberend F et al (2008) A cryptic vascular endothelial growth factor T-cell epitope: identification and characterization by mass spectrometry and T-cell assays. Cancer Res 68:2447–2454
278. Seward RJ, Drouin EE, Steere AC, Costello CE (2010) Peptides presented by HLA-DR molecules in synovia of patients with rheumatoid arthritis or antibiotic-refractory Lyme arthritis. Mol Cell Proteomics 10:M110.002477
279. Sahin U, Derhovanessian E, Miller M et al (2017) Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. Nature 547:222–226
280. Yadav M, Jhunjhunwala S, Phung QT et al (2014) Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. Nature 515:572–576
281. Gygi SP, Rist B, Griffin TJ et al (2002) Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. J Proteome Res 1:47–54

282. Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169
283. Cagney G, Emili A (2002) De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. Nat Biotechnol 20:163–170
284. Bozzacco L, Yu H, Zebroski HA et al (2011) Mass spectrometry analysis and quantitation of peptides presented on the MHC II molecules of mouse spleen dendritic cells. J Proteome Res 10:5016–5030
285. Hogan KT, Sutton JN, Chu KU et al (2005) Use of selected reaction monitoring mass spectrometry for the detection of specific MHC class I peptide antigens on A3 supertype family members. Cancer Immunol Immunother 54:359–371
286. Ishioka GY, Lamont AG, Thomson D et al (1992) MHC interaction and T cell recognition of carbohydrates and glycopeptides. J Immunol 148:2446–2451
287. Carbone FR, Gleeson PA (1997) Carbohydrates and antigen recognition by T cells. Glycobiology 7:725–730
288. Kastrup IB, Andersen MH, Elliott T, Haurum JS (2001) MHC-restricted T cell responses against posttranslationally modified peptide antigens. Adv Immunol 78:267–289
289. Kastrup IB, Stevanovic S, Arsequell G et al (2000) Lectin purified human class I MHC-derived peptides: evidence for presentation of glycopeptides in vivo. Tissue Antigens 56:129–135
290. Haurum JS, Arsequell G, Lellouch AC et al (1994) Recognition of carbohydrate by major histocompatibility complex class I-restricted, glycopeptide-specific cytotoxic T lymphocytes. J Exp Med 180:739–744
291. Haurum JS, Tan L, Arsequell G et al (1995) Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. Eur J Immunol 25:3270–3276
292. Haurum JS, Høier IB, Arsequell G et al (1999) Presentation of cytosolic glycosylated peptides by human class I major histocompatibility complex molecules in vivo. J Exp Med 190:145–150
293. Glithero A, Tormo J, Haurum JS et al (1999) Crystal structures of two H-2Db/glycopeptide complexes suggest a molecular basis for CTL cross-reactivity. Immunity 10:63–74
294. Zhang H, Li X-J, Martin DB, Aebersold R (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. Nat Biotechnol 21:660–666
295. Malaker SA, Ferracane MJ, Depontieu FR et al (2017) Identification and characterization of complex glycosylated peptides presented by the MHC class II processing pathway in melanoma. J Proteome Res 16:228–237
296. Morita D, Sugita M (2016) Lipopeptides: a novel antigen repertoire presented by major histocompatibility complex class I molecules. Immunology 149:139–145
297. Andersen MH, Bonfill JE, Neisig A et al (1999) Phosphorylated peptides can be transported by TAP molecules, presented by class I MHC molecules, and recognized by phosphopeptide-specific CTL. J Immunol 163:3812–3818
298. Meyer VS, Drews O, Günder M et al (2009) Identification of natural MHC class II presented phosphopeptides and tumor-derived MHC class I phospholigands. J Proteome Res 8:3666–3674
299. Zarling AL, Ficarro SB, White FM et al (2000) Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. J Exp Med 192:1755–1762
300. Mohammed F, Stones DH, Zarling AL et al (2017) The antigenic identity of human class I MHC phosphopeptides is critically dependent upon phosphorylation status. Oncotarget 8:54160–54172
301. Zarling AL, Polefrone JM, Evans AM et al (2006) Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc Natl Acad Sci U S A 103:14889–14894
302. Depontieu FR, Qian J, Zarling AL et al (2009) Identification of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. Proc Natl Acad Sci U S A 106:12073–12078
303. Li Y, Depontieu FR, Sidney J et al (2010) Structural basis for the presentation of tumor-associated MHC class II-restricted phosphopeptides to CD4+ T cells. J Mol Biol 399:596–603
304. Ferreira L, Sánchez-Juanes F, Munoz-Bellido JL, González-Buitrago JM (2011) Rapid method for direct identification of bacteria in urine and blood culture samples by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: intact cell vs.

extraction method. Clin Microbiol Infect 17:1007–1012

305. Welker M (2011) Proteomics for routine identification of microorganisms. Proteomics 11:3143–3153
306. Welker M, Moore ERB (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 34:2–11
307. Sogawa K, Watanabe M, Sato K et al (2012) Rapid identification of microorganisms by mass spectrometry: improved performance by incorporation of in-house spectral data into a commercial database. Anal Bioanal Chem 403:1811–1822
308. El-Bouri K, Johnston S, Rees E et al (2012) Comparison of bacterial identification by MALDI-TOF mass spectrometry and conventional diagnostic microbiology methods: agreement, speed and cost implications. Br J Biomed Sci 69:47–55
309. Ouedraogo R, Flaudrops C, Ben Amara A et al (2010) Global analysis of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One 5:e13691
310. Ouedraogo R, Daumas A, Ghigo E et al (2012) Whole-cell MALDI-TOF MS: a new tool to assess the multifaceted activation of macrophages. J Proteome 75:5523–5532
311. Portevin D, Pflüger V, Otieno P et al (2015) Quantitative whole-cell MALDI-TOF MS fingerprints distinguishes human monocyte sub-populations activated by distinct microbial ligands. BMC Biotechnol 15:24



Chapter 3

Classical Immunoproteomics: Serological Proteome Analysis (SERPA) for Antigen Identification

Kelly M. Fulton, Anna Ananchenko, Lawrence Wolfraim, Shannon Martin, and Susan M. Twine

Abstract

The study of the humoral immune response to infectious and chronic diseases is important for understanding the disease progression, identification of protective antigens, vaccine development, and discovery of biomarkers for early diagnosis. Proteomic approaches, including serological proteome analysis (SERPA), have been used to identify the repertoire of immunoreactive proteins in various diseases. In this chapter, we provide an outline of the SERPA approach, using the analysis of sera from mice vaccinated with a live attenuated tularemia vaccine as an example.

Key words SERPA, Serological proteome analysis, Western blotting, Immunoproteomics, Comparative immunoproteomics, *Francisella*, Tularemia, Live vaccine strain, Immunoreactivity, Antigens, Biomarkers

1 Introduction

The humoral immune response to various diseases has been studied for many years. Historically, this has been accomplished by methods such as agglutination [1], enzyme-linked immunosorbent assay (ELISA) [2], and one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) combined with Western blotting [3]. Early work was often unable to definitively identify the protein targets of the humoral immune response due to limitations of protein identification technologies. Since the early 2000s, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) combined with Western blotting has been exploited to separate protein antigens and permit improved resolution of proteins reactive with immune sera. This approach is also known as *se*rological *p*roteome *a*nalysis (SERPA). Combined with advances in genomics and mass spectrometry, SERPA has allowed the facile identification of individual immunoreactive proteins.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_3, © Springer Science+Business Media, LLC, part of Springer Nature 2019

The antigen used in these studies is usually a total proteome or subproteome fraction (e.g., membrane proteins) of target cell populations, such as, in vitro grown bacteria or cancerous cells from a tumor. 2D-PAGE is able to resolve many proteins in a complex mixture to single protein spots, retaining the native protein processing and posttranslational modifications. Subsequently, proteins are transferred to a nitrocellulose or polyvinylidene fluoride (PVDF) membrane and, probed with primary sera and then conjugated secondary antibody, as per traditional Western blotting. Proteins may be stained after transfer to a membrane, and the captured image can be used to align regions of immunoreactivity with areas of protein staining. Excising the identified immunoreactive proteins from a second protein-stained 2D-PAGE and subsequent digestion with trypsin allows identification of proteins using tandem mass spectrometry. This process is shown in flowchart format in Fig. 1. As with any experimental approach, 2D-Western blotting has well-documented disadvantages that are reviewed elsewhere [4]. Nevertheless, SERPA remains one of the most accessible immunoproteomics approaches and can be carried out in any laboratory equipped with protein electrophoresis and electroblotting equipment.

Antigen identification by SERPA is a versatile technique that can support research into a variety of diseases that involve the humoral immune response. In the last 20 years, a number of studies have employed SERPA for the identification of antigens in infectious diseases. For example, bacterial antigens identified by SERPA can be used to support vaccine development [5–8], or as immune correlates of protection and biomarkers [9] for disease diagnosis. Immunoproteomics has also been more recently used in the study of various cancers to identify potential biomarkers for early cancer diagnosis [10]. Biomarkers from patient immune sera, such as autoantibodies targeted at tumor associated antigens (TAA), provide the opportunity to detect cancer prior to the onset of symptoms [7]. This not only improves prognosis due to early diagnosis but also allows the monitoring of therapy [11]. In some cases, SERPA is used in combination with other immunoproteomics approaches such as SEREX [12] or immunohistochemical analysis [13]. Recent studies have used SERPA to identify serum autoantibodies as biomarkers for cancers, including colorectal cancer, prostate cancer, cholangiocarcinoma, lung cancer, and hepatocellular carcinoma, to name only a few [14–18]. A particularly interesting application of SERPA that bridges infectious diseases and cancer has been the identification of bacterial antigens from *Helicobacter pylori* that are predictive of the development of gastric cancer [19]. There have even been reports of SERPA being used to identify autoantigens involved in autoimmune disorders, such as multiple sclerosis (MS) [20], autoimmune polyendocrine syndrome type 1 (APS 1) [21], rheumatoid arthritis [22], and second-

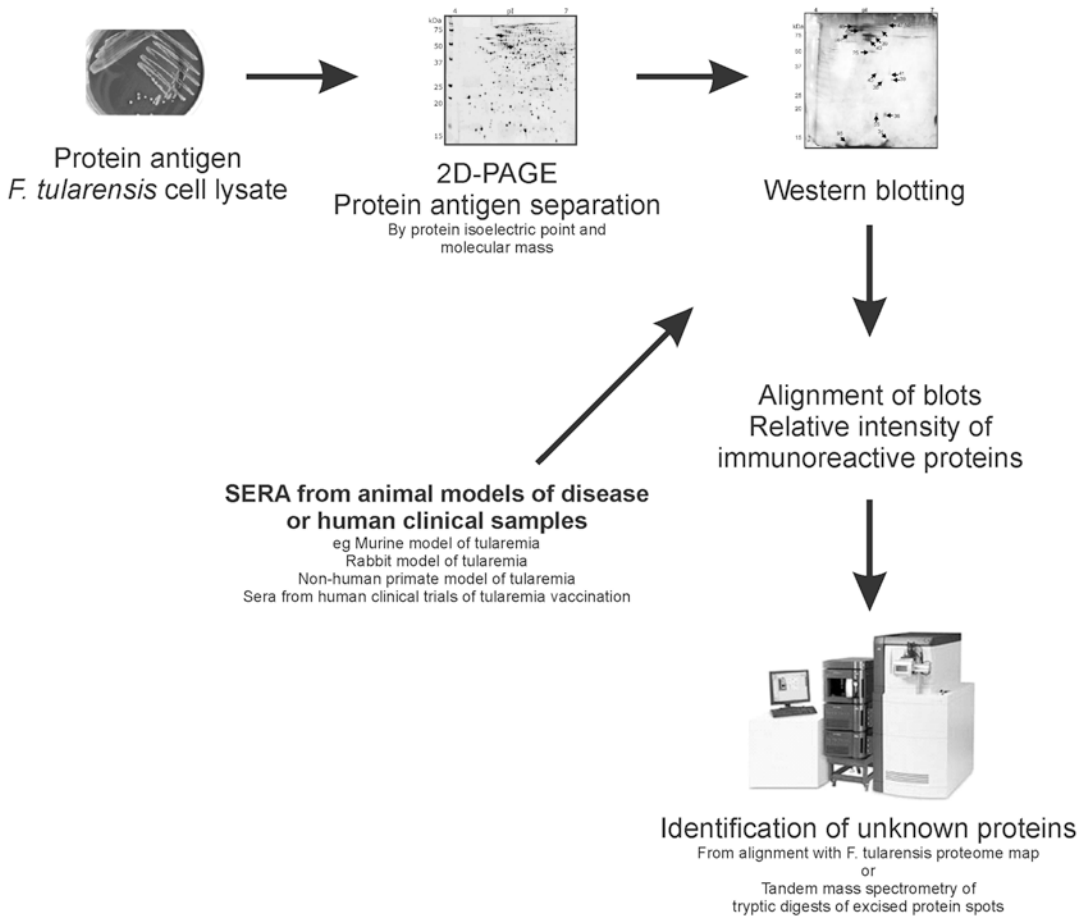


Fig. 1 Flowchart illustrating the approach used for SERPA. Protein antigens used for Western blotting were prepared from whole cell lysates of ***Francisella tularensis***. Proteins were then separated in two dimensions: by protein isoelectric point and by molecular mass. Resolved proteins were transferred to PVDF membranes by electroblotting and probed with sera from mice, rabbits, Non-human primates (NHPs), and humans. Western blots were aligned with protein-stained membranes of 2D-PAGE and immunoreactive proteins identified from a proteome map. Unknown proteins were identified using standard proteomics approaches, as described in [30]

ary Raynaud's [23]. The link between infectious diseases and autoimmune disorders can also be evaluated using SERPA, as has been done for *Campylobacter jejuni* which is associated with Guillain-Barré Syndrome (GBS) [24].

This chapter details a SERPA protocol that can be applied to many models of disease. In our laboratory, it was developed specifically for screening immune sera to support the development and characterization of a tularemia vaccine. Tularemia, a disease of humans and other mammals, is caused by the intracellular pathogen, *Francisella tularensis*. Although humans are an accidental host, *F. tularensis* is increasingly being isolated from infected humans in the United States and several European countries [25].

F. tularensis has gained significant attention in the post-9/11 era as one of six organisms designated as high priority agents by the US Center for Disease Control and Prevention that could be exploited as agents of bioterror (category A pathogens). It has been documented that *F. tularensis* was weaponized by both the United States and the former Soviet Union [26]. Combined, the low infectious dose and ease of dissemination of type A *F. tularensis* have made it a threat to both military personnel and civilians alike [27]. There is currently no licensed vaccine available in North America, although an attenuated type B strain, known as the live vaccine strain (LVS), has been used to vaccinate at-risk military personnel and laboratory workers. To further characterize the immune response induced by vaccination with LVS, we and others have exploited SERPA in an attempt to determine correlates of protection or markers of vaccination in animal models of disease and infected or vaccinated humans [28–43].

2 Materials

Prepare all solutions using deionized water with a resistance of ≥18 MΩ and analytical grade reagents (unless otherwise stated). Prepare and store all reagents at room temperature (unless indicated otherwise). Ensure appropriate safety precautions are followed and dispose of waste materials as per waste disposal guidelines. Sodium azide is not added to reagents.

2.1 Antigen Preparation

1. Cell lysis solution: 7 M urea, 2 M thiourea, 4% CHAPS, 1% DTT, 0.5% amidosulfobetaine-14 (ASB-14). Weigh 4.2 g of urea, 1.52 g of thiourea, 0.4 g of CHAPS, and 0.1 g of dithiothreitol (DTT), and add to 5.2 mL of water. Vortex vigorously. Then add 0.05 g of ASB-14 and mix with gentle inversion (*see* **Note 1**).
2. Cell lysis solution can be prepared ahead and divided into 1 mL aliquots. Store at −20 °C (*see* **Note 2**).

2.2 Isoelectric Focusing (IEF) Components

1. Immobilized pH gradient (IPG) strips (e.g., 17 cm pH 4–7 strips from Bio-Rad).
2. IPG solution: Add 20 μL of BioLyte® 3–10 stocks (Bio-Rad) to 180 μL of cell lysis solution and 10 μg of Orange G.
3. Disposable rehydration tray.
4. Light mineral oil.
5. Paper electrode wicks.
6. Isoelectric focusing cell, for example, PROTEAN® isoelectric focusing cell, including IEF tray from Bio-Rad.

2.3 Polyacrylamide Gel Electrophoresis (PAGE) Components

1. Resolving gel buffer: 1.5 M Tris(hydroxymethyl)aminomethane hydrochloride (Tris–HCl).
2. 30% Bis-acrylamide solution.
3. Sodium dodecyl sulfate (SDS): 10% stock solution. Weigh 0.1 g of SDS and add to 1 mL of water (*see* **Note 3**).
4. Ammonium persulfate (APS): 10% stock solution. Weigh 0.1 g of APS and add to 1 mL of water (*see* **Note 4**).
5. *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED).
6. SDS-PAGE running buffer: 1 × Tris/Glycine/SDS (TGS) running buffer diluted from 10 × stock (for example from Bio-Rad) with water (*see* **Note 5**).
7. Precision Plus Protein™ Dual Colour molecular weight marker (Bio-Rad) (*see* **Note 6**).
8. Agarose solution: Add 10 μg of Bromophenol blue to 50 mL ReadyPrep overlay agarose (*see* **Note 7**).
9. Equilibration base solution: 50 mM Tris–HCl, 6 M urea, 30% glycerol in water. Mix 0.68 mL of 1 M Tris–HCl, pH 8.8, 3.6 g of urea, and 3 g of glycerol, and then bring to a final volume of 10 mL with water. Vortex vigorously and divide into two 5 mL aliquots (*see* **Note 8**).
10. Reducing equilibration solution: 1% dithiothreitol (DTT), 2% SDS. Add 0.5 g of DTT to 5 mL of equilibration base solution. Vortex vigorously. Then add 0.1 g of SDS and mix gently by inversion (*see* **Note 8**).
11. Alkylating equilibration solution: 4% iodoacetamide (IODO), 2% SDS. Add 0.2 g of IODO to 5 mL of equilibration base solution. Vortex vigorously. Then add 0.1 g of SDS and mix gently by inversion (*see* **Notes 8** and **9**).
12. PROTEAN® II xi 2D cell system (Bio-Rad), including lower buffer chamber (buffer tank), cooling core/electrode assembly, cooling core gaskets, lid with electrical leads, outer and inner glass plates, gel casting stand, gel casting gaskets, sandwich clamps, IPG strip format combs, IPG strip format spacers, alignment card (*see* **Note 10**).
13. HC PowerPac™ power supply (Bio-Rad) (*see* **Note 11**).

2.4 Immunoblotting Components

1. Trans-Blot® semidry transfer cell (Bio-Rad).
2. HC PowerPac™ power supply (Bio-Rad) (*see* **Note 11**).
3. PVDF membrane.
4. Extra thick filter paper.
5. Phosphate buffered saline (PBS) and Tween (PBST): 8 g of sodium chloride, 0.2 g of potassium chloride, 1.44 g of disodium phosphate, 0.24 g of potassium dihydrogen phosphate, in 1 L of water, adjusted to pH 7.4) and 0.5 mL/L of Tween® 20.

6. PBST-M solution: 5% milk in PBST. Add 20 g of fat-free skim milk powder to 400 mL of PBST buffer (*see* **Note 12**).
7. Primary antibody and horseradish peroxidase (HRP) labeled secondary antibody (*see* **Note 13**).
8. Methanol: 100%.
9. Transfer buffer: 48 mM Tris, 39 mM glycine, 0.0375% SDS, in 20% methanol. Weigh 5.82 g of Tris, 2.93 g of glycine, and 0.375 g of SDS, and add to 200 mL of methanol. Bring to a final volume of 1 L with water. Pre-chill at 4 °C for 1 hour.
10. Fixing solution: 10% methanol, 7% acetic added. Add 20 mL of methanol and 14 mL of acetic acid to water for final volume of 200 mL. 200 mL is sufficient for one membrane.
11. SYPRO Ruby protein blot stain: 50 mL/membrane.
12. Enhanced chemiluminescence (ECL) detection kit (e.g., Bio-Rad Clarity™ Western ECL substrates): Prepared at a ratio of 1:1 for peroxide/enhancer reagents.
13. Kodak X-ray film.
14. Developer solution: Dilute 217 mL of concentrated Kodak developer stock solution with 783 mL water, as recommended by the manufacturer (*see* **Note 14**).
15. Fixer solution: Dilute 217 mL of concentrated Kodak fixer stock solution with 783 mL water, as recommended by the manufacturer (*see* **Note 14**).
16. Gel imager (e.g., Bio-Rad ChemiDoc™ imaging system with image analysis software such as Image Lab).

3 Methods

3.1 Sample Preparation

Unless otherwise specified, carry out all procedures at room temperature.

3.1.1 Serum

The serum used to illustrate the results of the current method was generated from animal models of the disease tularemia, caused by the bacterium *F. tularensis*. The description of how the serum was generated is described elsewhere [30]. The method is compatible with sera from humans or animals, as described in the accompanying notes.

3.1.2 Antigen Preparation

1. Thaw a vial of *F. tularensis* or O-antigen deficient strain *Δ wbtI* (used to avoid immunoreactivity of LPS) freezer stocks, prepared as described in [20]. Streak onto cysteine heart agar plates and grow overnight at 37 °C.

2. Scrape bacterial colonies from the agar plate using an inoculating loop or sterile slide, and transfer to a sterile microcentrifuge tube containing 1 mL of sterile deionized water.
3. Vortex the bacteria to resuspend, and then pellet by centrifugation at 22,000 × *g* (14,500 rpm in Eppendorf model) in a benchtop centrifuge.
4. Remove the supernatant, and replace with 1 mL of sterile deionized water, and repeat **step 3**.
5. Remove supernatant and add 12 times the pellet volume of cell lysis solution. Vortex until the solution clears (*see* **Note 15**).
6. Quantify the protein content of the cell lysates using a protein quantification assay, such as the Bradford assay [44] or similar assay, as per the kit instructions.

3.2 Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE)

1. Transfer 100 µg of protein antigen into a clean microfuge tube, and dilute to a total volume of 300 µL with cell lysis solution (*see* **Note 16**).
2. Add 15 µL Biolytes solution containing orange G dye (*see* **Note 16**).

3.2.1 Rehydration of IPG Strips

3. Shake the solution for 1 h at room temperature.
4. Centrifuge at 22,000 × *g* (14,500 rpm in Eppendorf model) in a benchtop centrifuge for 30 s to bring the solution to the bottom of the tube.
5. Pipette the solution into a rehydration tray, and overlay with one 17 cm IPG strip, pH 4–7, ensuring contact between exposed gel and solution. Incubate the strips for 1 h at room temperature (*see* **Note 17**).
6. Overlay each strip with 1.5 mL of mineral oil, and incubate overnight at room temperature (*see* **Note 18**).

3.2.2 First-Dimension Separation: Isoelectric Focusing (IEF)

1. Place paper wicks over the electrodes of the IEF focusing tray. Add 8 µL of water to each wick (*see* **Note 19**).
2. Remove the rehydrated IPG strips from the rehydration tray using forceps, and drain excess mineral oil. Lay strip onto IEF tray, gel side down, ensuring good contact with both electrodes. Overlay each strip with light mineral oil.
3. Place the tray containing strips into the IEF cell, and carry out IEF using the following steps at 20 °C: 200 V for 1 h, 500 V for 1 h, ramp to 5000 V over 5 h, focus to 80,000 Vh at 5000 V, and maintain at 500 V until a cumulative total of 95,000 Vh has been reached (approximately 24 h total) (*see* **Note 20**).

3.2.3 Second-Dimension Separation: PAGE

1. One day prior to commencing the SDS-PAGE separation portion of the protocol, prepare gels as follows.

2. Assemble large gel plates, as described by the manufacturer.
3. Prepare gel mix in a vacuum flask: 40 mL of 30% Bis/Acrylamide, 25 mL of 1.5 M Tris–HCl, pH 8.8, 33.5 mL of water, 1 mL of 10% SDS (*see* **Note 21**).
4. Degas the gel mix for 5 min.
5. Immediately before pouring the gels, add 400 μL of 10% APS and 40 μL of TEMED to 80 mL of gel mix.
6. Pipette the solution from **step 5** into assembled gels plates, taking care to avoid introducing air bubbles. Insert comb between the top of the two plates, and allow the gel solution to polymerize at room temperature for 1 h.
7. Transfer polymerized gels into an airtight bag containing 100 mL of water, and store at 4 °C until use the following day.
8. Immediately prior to running the SDS-PAGE separation, place focused strips into equilibration tray, gel side up.
9. Equilibrate each focused IEF strip with 2 mL of DTT solution at room temperature for 20 min (*see* **Note 22**).
10. Remove DTT solution, and equilibrate each IEF strip with 2 mL of IODO solution at room temperature for 20 min (*see* **Note 22**).
11. Heat agarose solution until melted and pipette 1 mL into each gel well. Place IEF strip into gel well, with gel side facing short plate, ensuring good contact with surface of SDS gel. Allow agarose to set for 10 min at 4 °C.
12. Assemble gels into holding clamps (as appropriate for the gel assembly used), and fill the upper loading chamber with SDS-PAGE running buffer (*see* **Note 23**).
13. Add 10 μL of molecular weight marker to the marker lane (*see* **Note 24**).
14. Add 1100 mL of SDS-PAGE running buffer to the tank before placing the plate assembly into tank (*see* **Note 25**).
15. Place plate assembly into tank slowly to avoid generating bubbles. Place lid on tank, and plug into PowerpPac™.
16. Run at 30 mA/gel for approximately 4.25 h, with water cooling (*see* **Note 26**).

3.3 Western Blotting and Protein Detection

3.3.1 Semidry Transfer

1. Incubate gel in 150 mL of transfer buffer for 30 min.
2. Rehydrate PVDF membrane for 10 s in methanol (*see* **Note 27**), and then rehydrate three filter papers (cut to size) and membrane in transfer buffer for 30 min (*see* **Note 28**).
3. Place one filter paper on platinum anode of semidry blotter, removing bubbles.

4. Place PVDF membrane on top of filter paper, and subsequently overlay with the SDS-gel (*see* **Note 29**).
5. Place two filter papers directly on top of the gel and attach second electrode, removing bubbles.
6. Transfer proteins from gel to PVDF membrane at 15 V for 1 h.
7. Remove PVDF membrane from blotting apparatus and air dry for 15 min (*see* **Note 30**).
8. Place the membrane protein side down in 200 mL of blot stain fixing solution for 15 min with gentle agitation.
9. Incubate membrane, protein side up in 200 mL of deionized water for 5 min. Repeat an additional three times.
10. Incubate membrane with 50 mL of SYPRO Ruby protein blot stain for 15 min with gentle agitation, protected from light.
11. Incubate membrane in deionized water for 1 min with gentle rotation. Repeat an additional two times.
12. Record image using protein gel scanner, such as ChemiDoc, with UV illumination for 5 s (*see* **Note 31**).
13. Air dry membrane for 15 min.
14. Rehydrate membrane for 5 s in methanol, and block with 50 mL of PBST-M solution overnight at 4 °C with agitation (*see* **Note 32**).

3.3.2 Probing and Fluorescence Detection of Immunoreactivity

1. Wash the membrane with 100 mL of PBST solution for 5 min, with gentle agitation. Repeat once. For all incubation steps, use gentle agitation (e.g., with an orbital shaker).
2. To prepare the primary mouse antiserum, dilute 50 μL of murine antiserum in 50 mL of PBST-M solution (1:1000 dilution) (*see* **Note 13**).
3. Remove PBST, and then add 50 mL of primary antiserum solution, and incubate for 1 h (*see* **Note 13**).
4. Remove the primary antiserum solution from membrane.
5. Wash the membrane with 100 mL of PBST, and incubate for 5 min.
6. Wash with 100 of mL fresh PBST solution for 30 min. Repeat this step an additional two times.
7. To prepare the secondary antibody solution, dilute 10 μL of goat anti-mouse IgG, HRP-labelled antibody in 50 mL PBST-M solution (1:5000 dilution) (*see* **Note 13**).
8. Remove PBST, and add 50 mL of secondary antibody solution, diluted appropriately, as denoted in this section (*see* **Note 13**). Incubate for 1 h.
9. Replace secondary antibody solution with 100 mL of PBST solution, and incubate for 5 min.

10. Remove PBST and add 100 mL of fresh PBST solution for 30 min. Repeat this step twice.
11. Remove PBST.
12. Mix Clarity ECL components at a 1:1 ratio, to a final volume of approximately 20 mL, and incubate membrane for 3 min (*see* **Note 33**).
13. Capture image of probed membranes using either film (Subheading 3.3.3) or a suitable fluorescence scanner (Subheading 3.3.4).

3.3.3 Developing Film Images and Alignment

Western-blotted membranes may be developed using film if an appropriate image scanner, capable of fluorescence detection, is not available (*see* **Note 34**).

1. Lay membrane between two transparency sheets.
2. Place membrane in autoradiography cassette.
3. In dark room, lay one sheet of Kodak Biomax Scientific imaging film on top of membrane. Close cassette and expose film for 1 min (*see* **Note 35**).
4. Incubate film in developer for 1 min.
5. Transfer to separate container of water and incubate film in water for 2 min.
6. Transfer film into fixer solution for a minimum of 5 min (*see* **Note 36**).
7. Transfer film to a second tray containing water for 2 min.
8. Air dry film.
9. Repeat **steps 1–8** for 30 s, 2 min, and 5 min film exposures (*see* **Note 35**).
10. Record image using protein gel scanner, with white light illumination for an appropriate length of time (*see* **Note 37**). Representative blots are shown in Fig. 2.
11. Scanned gel images can be aligned with the image of the protein-stained membrane or an equivalent protein-stained 2D-PAGE using image analysis software, such as PDQuest.
12. Protein identification can be carried out using now standard proteomics methods, as outlined in [30]. Figure 3 shows a representative silver-stained gel, indicating identified immunoreactive proteins (*see* **Note 38**) corresponding to the Western blots in Fig. 2.

3.3.4 Imaging Immunoreactivity Using ChemiDoc Gel Imaging System (Bio-Rad)

1. Using Image Lab software, set ChemiDoc™ scanner to default Western blot settings.

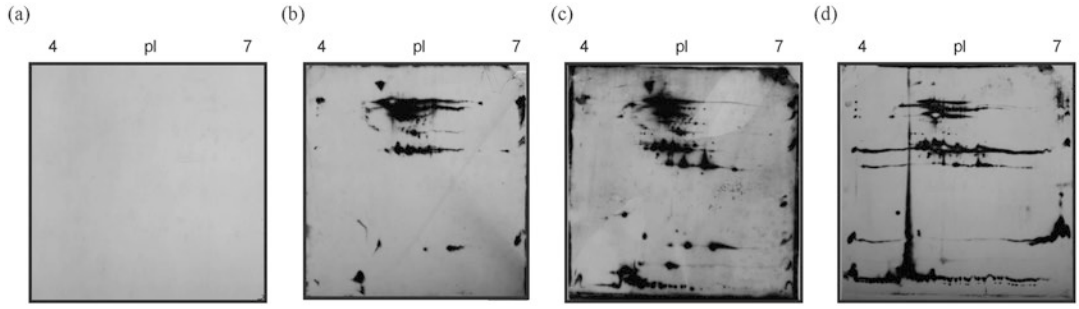


Fig. 2 2D Western blots probed with sera from mice vaccinated intranasally with ***F. tularensis*** LVS. Balb/c mice were immunized with lot 17 LVS or saline intranasally (~2.5 × 10^3 cfu per dose) and terminally bled. Shown are representative blots, probed with sera from individual mice (**a**) SHAM immunized, and LVS immunized mice (**b**) 4 weeks postvaccination, (**c**) 6 weeks postvaccination, and (**d**) 8 weeks postvaccination

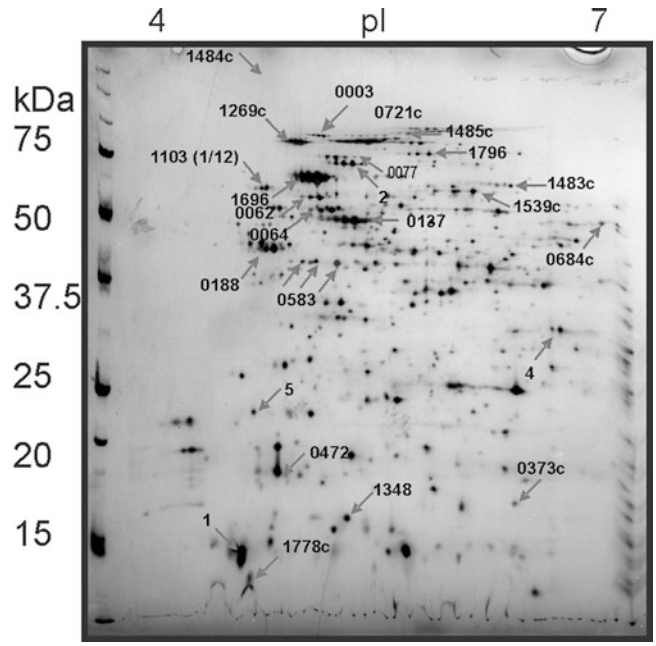


Fig. 3 2D-PAGE of total cell lysate of ***F. tularensis***. 100 μg of protein was separated in the pH range of 4–7. Proteins were visualized using non-fixing silver staining

2. Carefully remove membrane, and lay it directly on the glass surface of scanner, removing air bubbles.
3. Using scanner software (e.g., Image Lab), align and focus the camera on the membrane.
4. Scan membrane over a range of exposure times (*see* **Note 39**).
5. Scanned blot images are then aligned with protein-stained gels for subsequent protein identification, as outlined in Subheading 3.3.3, **steps 11** and **12**.

4 Notes

1. When adding water to dissolve the urea powder, note that the solution will initially be cold to touch. Mix frequently to dissolve completely.
2. Urea is used as a denaturant to increase the solubilization of proteins; however it is not stable. Do not heat urea containing solutions or store above 20 °C. In the presence of heat, urea breaks down to form isocyanate, which leads to carbamylation of proteins. Protein carbamylation will result in artifacts on 2D-PAGE. In addition, carbamylation of protein N-termini or lysine side chains can interfere with downstream enzymatic digestion and protein identification by mass spectrometry.
3. SDS is a respiratory, skin, and eye irritant. Weigh in fume hood. The stock solution may form crystalline precipitate in colder temperatures, so warm slightly to redissolve.
4. APS is prepared fresh for each use.
5. For better reproducibility of 2D-PAGE, we use a commercial 10× stock solution, purchased, for example, from Bio-Rad. Dilute the stock in deionized water, and stir slowly with magnetic stirrer to avoid formation of bubbles.
6. Prestained markers allow monitoring of protein transfer efficiency; at the end of the transfer, prestained markers will be visible on the membrane, and little or no marker will be visible in the gel.
7. The agarose solution can be divided into 5 mL aliquots, allowed to solidify, and stored at 4 °C. Prior to use, heat gently in water on a hot plate until liquefied.
8. The base solution of the reducing and alkylating equilibration solutions is the same. Therefore, first prepare 10 mL of equilibration base solution, and vortex vigorously as it will take time for the urea to completely dissolve. Once mixed, divide the equilibrating base solution into two 5 mL aliquots before adding DTT and IODO in Subheading 2.3, **steps 10** and **11**, respectively. Vortex both thoroughly before adding SDS. Do not vortex once SDS is added to avoid foaming; instead mix gently by inverting a few times.
9. IODO is light sensitive. The alkylating equilibration solution should therefore be made fresh and be kept in the dark.
10. 2D-PAGE equipment can be purchased as a complete system from many vendors, for example, the PROTEAN® II xi 2D cell system (excluding the power supply) from Bio-Rad. Individual components may also be purchased separately.

11. Vendors such as Bio-Rad offer multiple power supply options. It is important to consider the outputs required for gels and blotting. For example, with the PROTEAN® II xi 2D cell, the power supply must be capable of outputting a constant 30 mA/gel (or 60 mA/port). Many basic power supplies are sufficient for 2D-PAGE. However, even though protein transfer is performed with a constant 15 V regardless of the number of Trans-Blot® semidry transfer cells operating simultaneously, each cell will draw approximately 0.5 A. The output of many basic power supplies may be insufficient, and high current power packs, such as the Bio-Rad HC PowerPac™ or equivalent, are therefore recommended to accommodate up to four PROTEAN® II xi 2D cells or four Trans-Blot® semidry transfer cells.
12. We buy commercially available skimmed milk (e.g., Carnation) powder for this purpose. After solubilizing, filter using a Whatman filter to remove particulates.
13. Recommended primary antisera and secondary antibody dilutions.

    Mouse:

    Primary antisera: Dilute 50 µL mouse antisera in 50 mL of PBST-M solution (1:1000 dilution).

    Secondary antibody: Dilute 0.5 µL of goat anti-mouse IgG, HRP-labelled antibody (PerkinElmer Life and Analytical Sciences) in 50 mL of PBST-M solution (1:5,000 dilution).

    Rabbit:

    Primary antisera: Dilute 20 µL of rabbit antisera in 50 mL of PBST-M solution (1:2500 dilution).

    Secondary antibody: Dilute 10 µL of goat anti-rabbit IgG, HRP-labelled antibody (PerkinElmer Life and Analytical Sciences) in 50 mL of PBST-M solution (1:5000 dilution).

    Nonhuman primate (NHP):

    Primary antisera: Dilute 100 µL of monkey antisera in 50 mL PBST-M solution (1:500 dilution).

    Secondary antibody: Dilute 1 µL of goat anti-monkey IgG, HRP-labelled antibody (Fitzgerald Industries) in 50 mL of PBST-M solution (1:50,000 dilution).

    Human:

    Primary antisera: Dilute 100 µL of human antisera in 50 mL of PBST-M solution (1:500 µL dilution).

    Secondary antibody: Dilute 0.5 µL of goat anti-human IgG, HRP-labelled antibody (PerkinElmer Life and Analytical

Sciences) in 50 mL of PBST-M solution (1:100,000 dilution).

14. Based on the Kodak technical information sheets, the recommended replenishing rates are 20 mL of developer and 25 mL of fixer for each 20.3 cm × 25.4 cm film that is processed. When three times the original volume has been replenished, the solutions should be completely replaced. At minimum, developer and fixer solutions should be completely replaced every 3 months. However, when performing multiple experiments, solutions are generally replaced bi-weekly.
15. If cell lysis does not occur immediately, incubate at room temperature for 30 min with agitation. If solution has not cleared, add 10–20 μL of cell lysis solution, and continue until cells are lysed. Centrifuge to remove cell debris, and if dealing with pathogenic bacteria, plate 10% of the sample on a suitable agar medium to check for sterility before removing from biocontainment facility.
16. Volumes noted are for large gel format 17 cm IPG strips. Typically, these can be loaded with 80–600 μg of protein. We find for analytical purposes 100 μg gives the best resolution of gel spots. Other IPG strips can be purchased that are compatible with PAGE equipment from other manufacturers.
17. When overlaying strips, ensure that the rehydration tray is clean and dry. Remove plastic backing from the strip, with care not to damage the gel. Overlay strip onto rehydration solution carefully, using forceps to grip the plastic end of the strip. Take care not to create air bubbles as the strip is overlaid. There is a high degree of manual dexterity involved in doing this. We suggest practicing using the plastic backing from the strips to overlay on water spiked with orange G.
18. If you wish to avoid overlaying strips with mineral oil, an alternative solution is to fill empty wells of the rehydration tray with water and tightly seal the lid. The moist atmosphere prevents the strips from drying out. This is only applicable in cases where the number of wells permits. In addition, care must be taken not to allow water to overflow into wells containing the rehydrating IPG strips.
19. Moistening the IEF wicks helps remove salt from samples.
20. Different IEF programs are used for different IPG strip lengths, shown in Table 1.
21. Amounts indicated are sufficient for two large 12% gels. For resolution of proteins between 15 and 150 kDa, a 12% gel is usually appropriate. For better resolution of lower molecular weight proteins, a higher percentage acrylamide should be

Table 1
IEF programs for IPG strips of different lengths

IPG strip	IEF method	Target volt hours (VHs)	Approximate duration (h)
7 cm	30 min at 200 V 30 min at 500 V 2 h ramp to 6500 V 3 h focusing at 6500 V Maintain at 500 V	18,000–25,000	5–6
11 cm	45 min at 200 V 45 min at 500 V 4 h ramp to 4000 V 35,000 VH focusing at 4000 V Maintain at 500 V	45,000–55,000	20–24

used. Conversely, for higher molecular weight proteins, a lower percentage acrylamide should be used. Gels can be stored for 1–2 days at 4 °C.

22. Typically, strips are placed on rocker or to agitate gently during equilibration, at room temperature.
23. The seal between the gel plates and the cooling core gaskets is important to ensure containment of the SDS-PAGE running buffer. To ensure the seal is sufficient, we recommend filling the upper buffer chamber half way initially to check for leaking. If there are no leaks, continue filling the chamber. If there are leaks, pour out the running buffer and disassemble the gel/cooling core assembly. Moistening the gaskets with water sometimes helps; remount the gels into the holding clamps and test again. If leaking persists, applying vacuum grease to the gaskets can improve the seal but is difficult remove from the glass plates. If the leaking is chronic, the gaskets may need to be replaced.
24. To load molecular weight marker, we recommend using a pipette and gel loading tip. These tips have an elongated end which allows the marker to be loaded directly into marker well with no overflow. The molecular weight marker is always added after the SDS-PAGE running buffer to avoid displacing the marker out of the well.
25. Add running buffer to gel tank slowly, avoiding creation of bubbles. If bubbles are formed below the gel plates, use a spatula or similar device to move bubbles to one side of the gel tank.

26. At 30 mA/gel, the run time is usually 4.25 h. Monitor the gel and stop the run once the dye front has just migrated off of the gel.
27. SDS, present in denaturing gels and SDS-protein complexes, promotes protein elution from the gel during transfer. However, SDS hinders binding of proteins to the membrane. Methanol is added to the transfer buffer because it removes SDS from the SDS-protein complexes, improving protein-membrane binding. PVDF membranes are recommended over nitrocellulose for western blotting because they are more stable in the presence of organic solvents. Unlike nitrocellulose, PVDF membranes must be activated by a pretreatment with 100% methanol prior to equilibration in transfer buffer.
28. Shrinking or expanding of the gel or membrane during transfer will result in an apparent blurring of proteins transferred to the membrane. To avoid this, the gel must be adequately equilibrated in transfer buffer, and neither the gel nor the membrane must be permitted to dry out. The gel, membrane, and filter papers must also be prevented from drying to ensure proper conductivity during the transfer.
29. Proteins may begin to diffuse passively into the membrane immediately upon contact. Adjustment of the gel position once it has been laid should therefore be avoided to prevent transfer of the same protein to multiple spots on the membrane.
30. At all stages, it is important to limit the handling of the membrane. Forceps, rather than gloved hands, should be used. When handling the membrane, grasping only the edge or corner (even with forceps) will minimize the introduction of artifacts that may become visible when developing the film.
31. One gel can be reserved and silver stained in order to use as a reference proteome map. Staining of the PVDF membrane to visualize proteins gives a reasonable reference map for aligning developed blot image on X-ray film with protein-stained images, but blot staining is not as sensitive as silver staining of reference 2D-PAGE, so the silver-stained reference gel is also frequently used. The blot staining step with SYPRO Ruby blot stain is not essential. A non-fixing silver stain protocol that is compatible with potential downstream proteomics protocols is given here [45].
32. Blocking of blots is usually carried out overnight at 4 °C for convenience. It can also be carried out at room temperature for 1–2 h with similar results. Blocking and all subsequent incubation of the membrane are typically done on a rocker or orbital shaker to agitate gently.

33. Apply ECL kits solution immediately before exposing and developing film as the substrate will be used up and fluorescence signal will reduce over time.
34. If you have access to a scanner equipped with chemiluminescence, this can be a more rapid alternative to traditional X-ray film. Usually chemiluminescence works with HRP-fluorescence detection, and recommendations are usually provided by the manufacturer of the scanner.
35. When adapting this protocol, or working with a new set of sera, we generally use a range of exposure times, typically 10 s, 30 s, 1 min, 2 min, and 5 min. This range allows the optimal exposure time to be determined for a particular serum set. In addition, we routinely expose and develop films at 30 s, 1 min, and 5 min. Although comparisons across the dataset were always made using 1-min exposure, shorter or longer exposures sometimes helped in alignment of immunoreactive areas with protein-stained gels, for example, where a shorter exposure helps separate an intensely immunoreactive region into several distinct areas.
36. The X-ray films are coated in silver halides which are converted to metallic silver during developing to produce an image. In addition to making the image permanent, the fixing step also dissolves unconverted silver halides from the rest of the film. Since these silver halides are light sensitive, residual silver halides, as a result of insufficient fixing time, will result in discoloration of the film once it is removed from the dark room. Therefore a minimum of 5 min should be used for the fixing step.
37. Alignment can be carried out using open source or commercially available gel alignment software, for example, PDQuest (Bio-Rad). This is highly recommended if measurement of relative intensity of each immunoreactive area by densitometry is required. If appropriate software is not available, then a scanned image of an equivalent silver-stained 2D-PAGE can be aligned manually with the developed blot.
38. Protein identification, using standard proteomics approaches can be carried out in house, or outsourced to a third party service provided.
39. Commencing with default software settings is helpful to find the optimal parameters. We also suggest manually recording exposure times at several time points. This can help with visualizing protein spots with poor immunoreactivity (longer exposure times) or delineating multiple clusters of protein spots (shorter exposure times). Both help with the alignment of blots with protein-stained gel images. We also note that any

comparison of immunoreactivity of intensity cannot be absolute quantification and is always carried out between blots with the same exposure time.

Acknowledgments

This work was supported by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, Contract No. HHSN266200500041C, and in part by the National Research Council, Canada. The authors thank Dr Freyja Lynn, Vicki Pierson, Kristin Debord, and Patrick Sanz (National Institutes of Allergy and Infectious Diseases); Dr J. Wayne Conlan, Luc Tessier, and Marianne Savicky (National Research Council, Canada); and Gretchen Stup (DynPort Vaccine Company) for their contributions throughout this work.

References

1. Engelfried JJ, Spear F (1966) Modified agglutination test for Pasteurella tularensis. Appl Microbiol 14:267–270
2. Carlsson HE, Lindberg AA, Lindberg G et al (1979) Enzyme-linked immunosorbent assay for immunological diagnosis of human tularemia. J Clin Microbiol 10:615–621
3. Bevanger L, Maeland JA, Naess AI (1988) Agglutinins and antibodies to Francisella tularensis outer membrane antigens in the early diagnosis of disease during an outbreak of tularemia. J Clin Microbiol 26:433–437
4. Kilmury SLN, Twine SM (2010) The Francisella tularensis proteome and its recognition by antibodies. Front Microbiol 1:143
5. Vytvytska O, Nagy E, Blüggel M et al (2002) Identification of vaccine candidate antigens of *Staphylococcus aureus* by serological proteome analysis. Proteomics 2:580–590
6. Klade CS (2002) Proteomics approaches towards antigen discovery and vaccine development. Curr Opin Mol Ther 4:216–223
7. Pizarro-Guajardo M, Ravanal MC, Paez MD et al (2018) Identification of *Clostridium difficile* immunoreactive spore proteins of the epidemic strain R20291. Proteomics Clin Appl 12:e1700182
8. Pang H-Y, Li Y, Wu Z-H et al (2010) Immunoproteomic analysis and identification of novel immunogenic proteins from Vibrio harveyi. J Appl Microbiol 109:1800–1809
9. Li H, Ye M-Z, Peng B et al (2010) Immunoproteomic identification of polyvalent vaccine candidates from *Vibrio parahaemolyticus* outer membrane proteins. J Proteome Res 9:2573–2583
10. Dai L, Li J, Tsay J-CJ et al (2017) Identification of autoantibodies to ECH1 and HNRNPA2B1 as potential biomarkers in the early detection of lung cancer. Oncoimmunology 6:e1310359
11. Parida SK, Kaufmann SHE (2010) The quest for biomarkers in tuberculosis. Drug Discov Today 15:148–157
12. Tan HT, Low J, Lim SG, Chung MCM (2009) Serum autoantibodies as biomarkers for early cancer detection. FEBS J 276:6880–6904
13. Desmetz C, Mange A, Maudelonde T, Solassol J (2011) Autoantibody signatures: progress and perspectives for early cancer detection. J Cell Mol Med 15:2013–2024
14. Forgber M, Trefzer U, Sterry W, Walden P (2009) Proteome serological determination of tumor-associated antigens in melanoma. PLoS One 4:e5199
15. Gao H, Zheng Z, Mao Y et al (2014) Identification of tumor antigens that elicit a humoral immune response in the sera of Chinese esophageal squamous cell carcinoma patients by modified serological proteome analysis. Cancer Lett 344:54–61
16. Klein-Scory S, Kübler S, Diehl H et al (2010) Immunoscreening of the extracellular proteome of colorectal cancer cells. BMC Cancer 10:70
17. Mustafa MZ, Nguyen VH, Le Naour F et al (2016) Autoantibody signatures defined by serological proteome analysis in sera from

patients with cholangiocarcinoma. J Transl Med 14:17
18. Dai L, Qu Y, Li J et al (2017) Serological proteome analysis approach-based identification of ENO1 as a tumor-associated antigen and its autoantibody could enhance the sensitivity of CEA and CYFRA 21-1 in the detection of non-small cell lung cancer. Oncotarget 8:36664–36673
19. Krah A, Miehlke S, Pleissner K-P et al (2004) Identification of candidate antigens for serologic detection of Helicobacter pylori-infected patients with gastric carcinoma. Int J Cancer 108:456–463
20. Colomba P, Fontana S, Salemi G et al (2014) Identification of biomarkers in cerebrospinal fluid and serum of multiple sclerosis patients by immunoproteomics approach. Int J Mol Sci 15:23269–23282
21. Grandjean M, Sermeus A, Branders S et al (2013) Hypoxia integration in the serological proteome analysis unmasks tumor antigens and fosters the identification of anti-phospho-eEF2 antibodies as potential cancer biomarkers. PLoS One 8:e76508
22. Biswas S, Sharma S, Saroha A et al (2013) Identification of novel autoantigen in the synovial fluid of rheumatoid arthritis patients using an immunoproteomics approach. PLoS One 8:e56246
23. Yang L, Fujimoto M, Murota H et al (2015) Proteomic identification of heterogeneous nuclear ribonucleoprotein K as a novel cold-associated autoantigen in patients with secondary Raynaud's phenomenon. Rheumatology 54:349–358
24. Loshaj-Shala A, Colzani M, Brezovska K et al (2018) Immunoproteomic identification of antigenic candidate Campylobacter jejuni and human peripheral nerve proteins involved in Guillain-Barré syndrome. J Neuroimmunol 317:77–83
25. Eliasson H, Broman T, Forsman M, Bäck E (2006) Tularemia: current epidemiology and disease management. Infect Dis Clin N Am 20:289–311, ix
26. Zilinskas RA (2017) A brief history of biological weapons programmes and the use of animal pathogens as biological warfare agents. Rev Sci Tech 36:415–422
27. Dennis DT, Inglesby TV, Henderson DA et al (2001) Tularemia as a biological weapon: medical and public health management. JAMA 285:2763–2773
28. Twine SM, Petit MD, Shen H et al (2006) Immunoproteomic analysis of the murine antibody response to successful and failed immunization with live anti-Francisella vaccines. Biochem Biophys Res Commun 346:999–1008
29. Twine SM, Mykytczuk NCS, Petit MD et al (2006) In vivo proteomic analysis of the intracellular bacterial pathogen, Francisella tularensis, isolated from mouse spleen. Biochem Biophys Res Commun 345:1621–1633
30. Twine SM, Petit MD, Fulton KM et al (2010) Immunoproteomics analysis of the murine antibody response to vaccination with an improved Francisella tularensis live vaccine strain (LVS). PLoS One 5:e10000
31. Thomas RM, Twine SM, Fulton KM et al (2011) Glycosylation of DsbA in Francisella tularensis subspecies tularensis. J Bacteriol. https://doi.org/10.1128/JB.00438-11
32. Havlasová J, Hernychová L, Halada P et al (2002) Mapping of immunoreactive antigens of Francisella tularensis live vaccine strain. Proteomics 2:857–867
33. Havlasová J, Hernychová L, Brychta M et al (2005) Proteomic analysis of anti-Francisella tularensis LVS antibody response in murine model of tularemia. Proteomics 5:2090–2103
34. Janovská S, Pávková I, Reichelová M et al (2007) Proteomic analysis of antibody response in a case of laboratory-acquired infection with Francisella tularensis subsp. tularensis. Folia Microbiol 52:194–198
35. Janovská S, Pávková I, Hubálek M et al (2007) Identification of immunoreactive antigens in membrane proteins enriched fraction from Francisella tularensis LVS. Immunol Lett 108:151–159
36. Huntley JF, Conley PG, Hagman KE, Norgard MV (2007) Characterization of Francisella tularensis outer membrane proteins. J Bacteriol 189:561–574
37. Huntley JF, Conley PG, Rasko DA et al (2008) Native outer membrane proteins protect mice against pulmonary challenge with virulent type A Francisella tularensis. Infect Immun 76:3664–3671
38. Straskova A, Spidlova P, Mou S et al (2015) Francisella tularensis type B ΔdsbA mutant protects against type A strain and induces strong inflammatory cytokine and Th1-like antibody response in vivo. Pathog Dis 73:ftv058
39. Post DMB, Slütter B, Schilling B et al (2017) Characterization of inner and outer membrane proteins from Francisella tularensis strains LVS and Schu S4 and identification of potential subunit vaccine candidates. MBio 8. https://doi.org/10.1128/mBio.01592-17

40. Njeru J, Tomaso H, Mertens K et al (2017) Serological evidence of Francisella tularensis in febrile patients seeking treatment at remote hospitals, northeastern Kenya, 2014–2015. New Microbes New Infect 19:62–66
41. Zákutná Ľ, Dorko E, Rimárová K, Kizeková M (2015) Pilot cross-sectional study of three zoonoses (Lyme disease, Tularaemia, Leptospirosis) among healthy blood donors in eastern Slovakia. Cent Eur J Public Health 23:100–106
42. Chandler JC, Sutherland MD, Harton MR et al (2015) Francisella tularensis LVS surface and membrane proteins as targets of effective post-exposure immunization for tularemia. J Proteome Res 14:664–675
43. Lu Z, Perkins HM, Sharon J (2014) Antibodies to both terminal and internal B-cell epitopes of Francisella tularensis O-polysaccharide produced by patients with tularemia. Clin Vaccine Immunol 21:227–233
44. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248–254
45. Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis 8:93–99



Chapter 4

Profiling of Cytokine and Chemokine Responses Using Multiplex Bead Array Technology

Greg Harris and Wangxue Chen

Abstract

Multiplex bead array technology expands upon the principles of the enzyme-linked immunosorbent assay by allowing the simultaneous quantification of a large number of cytokines and chemokines within a single sample. This allows researchers more freedom and opportunities to investigate complex immune responses both in vivo and in vitro. Here we describe and update the detailed assay protocol and technical tips for simultaneous quantification of multiple cytokines and chemokines in mouse biological fluids such as sera, bronchoalveolar lavage fluid, tissue homogenate supernatant, and tissue culture supernatant, using a multiplex bead array assay.

Key words Multiplex assay, Cytokine, Chemokine, Bead array, Immunoassay, Luminex

1 Introduction

Multiplex bead array assays (MBAAs) are a highly efficient alternative to traditional enzyme-linked immunosorbent assays (ELISA) for quantification of cytokines and chemokines in various types of biological samples [1, 2]. In a traditional ELISA, usually one analyte is studied at a time, and the number of possible analytes is limited by sample volume. Typically, 50–100 μL of sample at a working dilution is needed for a single analyte ELISA; therefore, only a few ELISA assays can be carried out when the sample has an inherently small volume (such as mouse sera and human cerebrospinal fluid). MBAAs allow for the simultaneous quantification of up to 50, 100, or 500 analytes (depending on the system used) in the same sample, and generally no more than 50 μL of sample is required for the entire panel.

Commercial MBAA reagents became first available in the late 1990s, developed by Luminex® (Austin, TX) under the name FlowMetrix™ [3–6], now known as xMAP® technology (multi-analyte profiling) (*see* **Note 1**). Initially, the assays were run and analyzed on a standard flow cytometer equipped with specialized

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_4, © Springer Science+Business Media, LLC, part of Springer Nature 2019

software. Luminex® later developed the Luminex 100®, a flow cytometer specifically equipped to acquire and analyze microspheres in an MBAA format. These units went into common use in the early 2000s [7, 8].

Since its inception, MBAA technology has been evaluated extensively [9–13] and has been cited in thousands of scientific publications [14]. MBAA have broad applications within the field of life science research, including cytokine and chemokine responses, assays for cell signaling molecules, cancer markers, metabolic markers, antibody isotyping, as well as pathogen detection [2, 15–19]. More recently MBAA are being used in PCR applications, particularly for pathogen detection in environmental and clinical diagnostic applications [20–23].

The principle of MBAA is similar to that of ELISA [24], except that capture antibodies are covalently coupled with the surface of polystyrene beads instead of the bottom of a 96-well plate (Fig. 1). Each well contains a mixture of bead types, each type coupled with a different capture antibody. The bead conjugates are distinguished from one another by the fluorescence intensity ratio of two or three different fluorescent dyes embedded within the bead. Sample incubation is similar to ELISA, as well as the detection steps, which use biotinylated detection antibodies and a streptavidin-phycoerythrin (PE) conjugate reporter. Sample acquisition is performed using a specialized flow cytometer or fluorescent imager, which perform a series of analyses similar to flow cytometric analysis of cell targets. Individual beads are first gated to remove doublets (i.e., beads that have adhered to one another) from analysis, then the bead type is determined based on the ratio of the internal fluorescent dyes, and finally the fluorescence intensity of the reporter is determined for each individual bead. This last number correlates with the concentration of a given analyte in solution, similar to the optical density readings in an ELISA assay. A standard curve is generated to calculate the final analyte concentration based on the median fluorescence intensity of the bead.

At present three types of analyzer are commercially available from Luminex®. The LS200® is the second version of the LS100® and is functionally similar to a benchtop flow cytometer, using a fluidics-based sample transport system and laser-based detection methods. MAGPIX® also uses a fluidics-based sample transport system, but instead of laser detection, it uses a magnetic plate to capture all beads simultaneously in a grid, light-emitting diode (LED) lights to illuminate the chamber and a charge-coupled device (CCD) image sensor system to capture a digital image of the beads and determine the fluorescence intensity of each bead. MAGPIX® is more compact and affordable than the LS200® system but currently only allows for analysis of 50 analytes at a time; this is not necessarily a limitation, as the number of commercially available analytes from any given supplier is typically less than 50.

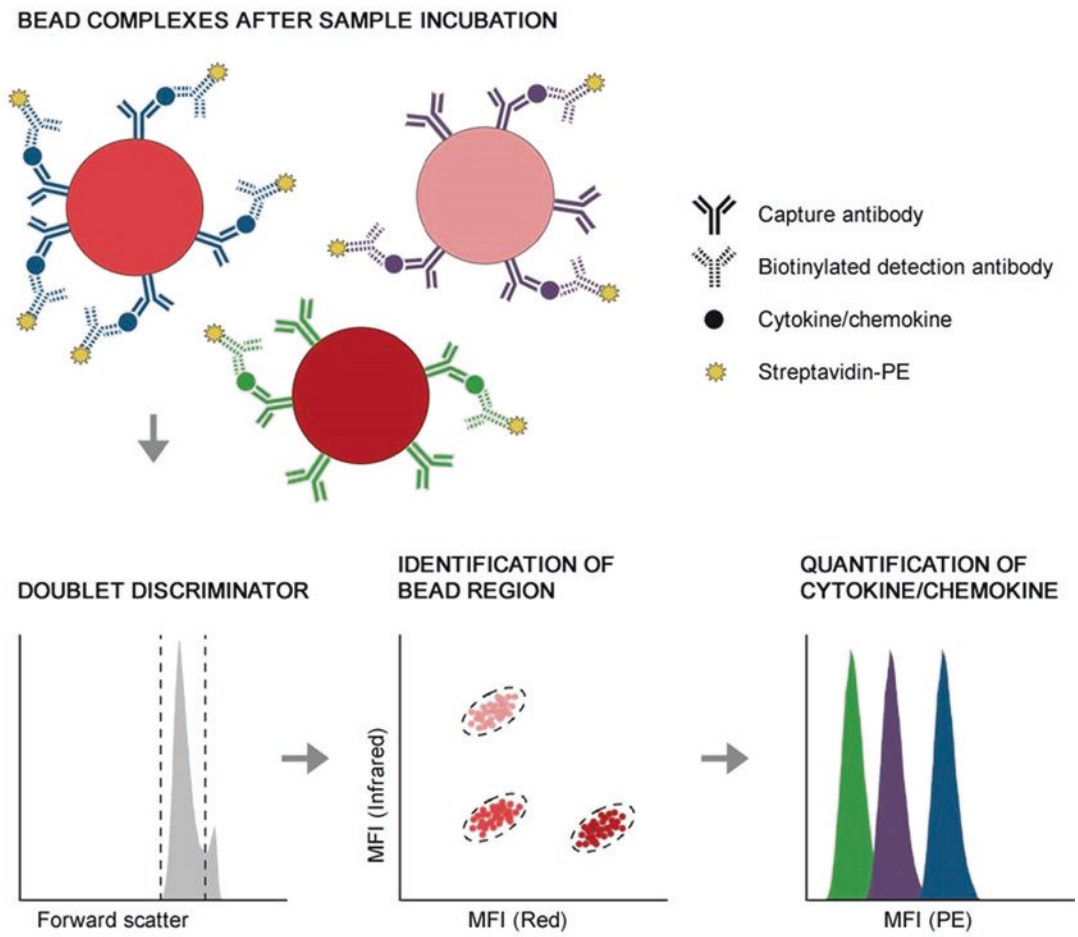


Fig. 1 The principle of multiplex bead array assay (MBAA). Monoclonal antibodies are covalently coupled to the surface of beads so that each bead type has a unique type of monoclonal antibody attached to its surface. Beads are eventually distinguished from one another based on the relative amounts of red and infrared dyes embedded within. Beads are mixed together, and the cocktail is added to each well of the assay, followed by incubation with samples, then with a biotinylated detection antibody, and finally with streptavidin-PE as the reporter. Bead complexes are passed through the fluidics system of a specialized flow cytometer. Beads are gated on their size and doublets are screened out. Beads are then selected based on predetermined gates defined by the ratio of red and infrared dyes. Finally, a reporter laser and photomultiplier determine the fluorescence intensity of the reporter color (PE) for each bead. When extrapolated on a standard curve, the fluorescence intensity correlates directly to the concentration of the specified analyte (i.e., cytokine or chemokine) in the sample

More recently, Luminex® has developed FLEXMAP 3D® for high-throughput sample analysis. FLEXMAP 3D® escalates the capabilities of the LS200® by combining three dyes within the microspheres to allow a multiplex number as high as 500 and by permitting the use of both 96- and 384-well plates.

MBAA reagents for common cytokines, chemokines, and immunological signaling pathways are available in pre-assembled commercial kits (*see* Table 1), which are usually highly customizable

Table 1
Common vendors of Luminex®-compatible MBAA kits (*see* Note 1)

Product line/company	Assay types	Species range	Website
Bio-Plex/Bio-Rad (Hercules CA)	Acute phase Angiogenesis Cancer Cell signaling Cytokine/chemokine Diabetes Isotyping	Human, mouse, canine and rat	www.bio-rad.com
Fujirebio (Belgium/Sweden)	Alzheimer's	Human	www.fujirebio-europe.com
Milliplex/Millipore (Burlington MA)	Adipokine Bone metabolism Cancer Cardiovascular Cell signaling Cellular metabolism Cytokine/chemokine Endocrine Isotyping Skin Toxicity	Human, mouse, rat, nonhuman primate, dog, guinea pig, hamster, pig, and rabbit	www.millipore.com
Myriad RBM (Austin TX)	Autoimmune and arthritis Bone disease Cancer Cardiovascular Cytokine/chemokine Diabetes and metabolic markers Endocrine Gastrointestinal Kidney toxicity Neurological disease	Human	www.myriadrbm.com
OriGene (Rockville MD)	Genotyping Cancer Transcription factors/nuclear receptors	Human	www.origene.com

(continued)

Table 1
(continued)

Product line/company	Assay types	Species range	Website
Procarta, Novex, Quantigene/Thermo Fisher (Santa Clara CA)	Cell signaling Cytokine/chemokine Gene expression profiling Transcription factors/nuclear receptors	Human, mouse, rat, nonhuman primate, dog, and pig	www.thermofisher.com
R&D Systems (Minneapolis MN)	Angiogenesis Cardiovascular Cytokine/chemokine Kidney toxicity MMP Obesity	Human, mouse, and rat	www.rndsystems.com

for a wide range of analytes. In addition, there is a vast array of non-immunological applications for MBAA analysis of various biomarkers implicated in human health. Most vendor websites offer user-friendly interface for finding and/or building kits of interests. Availability of analytes varies from one vendor to the next, but the majority of commercially available kits are intended for use on an LS200® or MAGPIX® analyzer, so there is a cross vendor conformity in the general principles of the assay, particularly in the beads used. Additionally, MBAA users have the option of creating custom bead assays by purchasing uncoated MAGPLEX™ beads and an xMAP® Antibody Coupling Kit from Luminex®. The antibody coupling process involves a small number of steps and about 4 h of total time. Assay optimization (standards, detection antibodies) must be performed by the user.

As with ELISA, the individual steps in the assay procedure are optimized for each commercial kit, based on the manufacturer's protocols. Furthermore, there are various system and software options for data acquisition and analysis. Here we outline the steps of an MBAA using a Milliplex® Mouse Cytokine/Chemokine kit (Millipore®, Burlington MA), which is compatible with an LS200® or a MAGPIX® analyzer for data acquisition. While the majority of these steps apply to both analyzers, some differences exist and these are indicated throughout.

2 Materials

1. Laboratory rocker or shaker (e.g., Lab Line Instruments Model #4625).
2. Aluminum foil.
3. Vortex mixer.
4. Sonicator (e.g., Branson® Ultrasonic Cleaner model #B200).
5. Handheld magnetic separation block (e.g., Millipore®), or automatic plate washer for magnetic beads (e.g., Bio-Tek ELx405, Millipore®).
6. LS200® or MAGPIX® analyzer.
7. Analysis software (e.g., Luminex xPONENT® software v4.2 or Millipore Analyst™ software v3.5.5.0).
8. Milliplex Mouse Cytokine/Chemokine custom kit (e.g., Millipore®).
9. Calibration/verification beads (Luminex®) (*see* **Note 2**).
10. Phosphate Buffered Saline (PBS) (LS200® only).
11. Ethanol: 70% (v/v).
12. Deionized water.
13. Sheath fluid (Luminex®).
14. Bleach: 10% (v/v) in water.
15. 0.1 M NaOH (MAGPIX® only).

3 Methods

All steps are to be carried out at room temperature (20–25 °C) unless otherwise stated. Ensure that all procedures after **step 4** are carried out under low light conditions to preserve fluorescence intensity of beads and the streptavidin-PE conjugate.

3.1 Setup and Incubation

1. Acclimate kit components to room temperature before use (*see* **Note 3**).
2. Presoak plate with 200 µL wash buffer, and place on a plate shaker or rotator at a medium-high speed (*see* **Note 4**) for 10 min.
3. Dilute lyophilized standards and quality control samples with 250 µL deionized H_2O, and mix gently. Leave at room temperature for at least 5 min, and then transfer to polypropylene microcentrifuge tubes.
4. Prepare standard curve dilutions in polypropylene microcentrifuge tubes, as outlined in the kit instructions.

5. Mix beads if necessary (*see* **Note 5**) by adding 60 µL of each bead type to the mixing bottle provided with the kit, then adding the appropriate volume of assay buffer for a total volume of 3 mL. Sonicate the bead cocktail for 30 s in an ultrasonic cleaner, and vortex at high speed for 60 s (*see* **Note 6**). Unused diluted beads can be stored at 4 °C for up to 1 month (*see* **Note 7**) but should be sonicated and vortexed each time before use.
6. Remove wash buffer by flicking it into a sink, gently blot the plate dry, and add 25 µL of sample diluent or standard diluent to appropriate wells (*see* **Note 8**).
7. Add 25 µL of standards, controls, and samples (*see* **Note 9**) to appropriate wells, in duplicate.
8. Add 25 µL of bead cocktail to all wells.
9. Seal plate with a plastic plate sealer, protect plate from light by wrapping or covering with aluminum foil, and secure on plate shaker or rocker. Shake at a medium-high speed for 2 h at room temperature, or overnight at 4 °C (*see* **Note 10**).
10. Use a manual or automatic plate washer to remove well contents, and wash twice with 200 µL of wash buffer.

    Manual washing (*see* **Note 11**): Affix the plate to the handheld separation block, and allow beads to settle at the bottom of wells for 60 s. Flick off liquid into a sink or similar receptacle. Typically this will involve three sharp downward motions to remove all the liquid and remaining droplets. With the plate still locked in the separation block, gently blot the surface of the plate with a paper towel to absorb any residual liquid. Do not tap the plate on paper towel as in an ELISA assay; despite the fact that beads are magnetically bound to the plate bottom, tapping can dislodge them. Remove plate from the separation block, add 200 µL of wash buffer to all wells, seal the plate, and return it to the rocker or shaker. Rock/shake the plate for 30 s at the same speed used for incubation steps, then reaffix plate to separation block, and repeat wash steps one more time (total of two washes).

    Automatic washing: Plates can also be washed with an automatic plate washer intended for magnetic beads. Follow plate washer instructions using the wash conditions laid out in the assay kit manual (typically 200 µL/well, two washes).
11. Add 25 µL detection antibody cocktail (provided premixed in the kit). Reseal plate, protect from light, and secure on plate shaker or rocker. Shake at a medium-high speed for 1 h at room temperature (*see* **Note 12**).
12. Do *not* perform a wash step at this point. Add 25 µL streptavidin-PE conjugate to all wells. Reseal plate, protect from light,

and secure on plate shaker or rocker. Shake at a medium-high speed for 30 min at room temperature.

13. Wash the wells twice as in **step 10**.
14. Resuspend beads in 150 μL of PBS or sheath fluid, reseal plate, protect from light, and secure on plate shaker or rocker. Shake at a medium-high speed for 5 min at room temperature, and then acquire samples (*see* **Note 13**).

3.2 Data Acquisition

Setup and operation of the LS200® and MAGPIX® analyzers is similar to that of a flow cytometer. Please refer the user manual provided with the analyzer for step-by-step detailed instructions. The following is a basic overview of the steps involved and the estimated time required:

1. System setup (30 min): Warming up lasers (LS200® only), priming fluidics with sheath fluid, cleaning with 70% ethanol, and washing (cleaning and priming can be performed while lasers are warming up).
2. Assay settings and plate layout (10–20 min): Each type of commercial kit has its own optimal settings (e.g., number of beads to be acquired, bead size gates, sample volume collected) which need to be inputted by the user. Also, bead ID data (i.e., which analyte is associated with which bead type) must be entered by the user. Finally, plate information is entered, including sample IDs, and information about standard curve dilutions and controls.
3. Calibration/verification (5–10 min) (*see* **Note 2**): The analyzer requires calibration and verification every time it is turned on, as well as once a month (LS200®) or once a week (MAGPIX®) if left on during that time. This involves running calibration beads, to calibrate the classification channels (which determine bead type) and reporter channels (which determine fluorescence intensity of PE). Control beads are then run in order to verify calibration.
4. Data acquisition (45–90 min): The plate is loaded into the analyzer, and once the data acquisition starts, the entire process is automated (*see* **Note 14**). It is recommended that the user does not leave the analyzer unattended for long periods of time, as warning alarms may sound during the acquisition process which demand the user's attention (i.e., low sheath fluid levels, sample error, fluidics error). Total acquisition time can vary from one assay to the next but generally takes no more than 1 min per well (*see* **Note 15**).
5. System shutdown (15 min): The fluidics systems are flushed with cleaning and decontaminating solutions (70% ethanol,

10% bleach, or 0.1 M NaOH, depending on the system used) and flushed thoroughly with deionized water to remove any remaining salts. The system is then shut down, or left on if it will be used in the near future (the LS200® lasers power down automatically if the system is left idle for a few hours).

3.3 Data Analysis

Multiplex analysis software will calculate the best-fit standard curve formulas and then determine the concentration of each analyte based on the MFI of its associated bead type in each well. The user can also select the best fit formula, if needed. The output is typically the average concentration of replicate wells, average MFI and % CV of replicates, minimum and maximum detection limits, and R^2 values of each standard curve. A full report of every well's raw data (pre-analysis) is also available in a text file table which displays MFI and total bead count for each bead type in each well. Data may also be analyzed manually, by inputting the MFI of each bead type in every well (instead of the optical density, as with ELISA) and using an analysis software of their choice to calculate the standard curve formulas and determine the concentration of each analyte. Because this must be performed manually for each analyte in the panel, it is not recommended due to time considerations; however it may be necessary at the user's discretion.

Luminex xPONENT™ is the acquisition/analysis software bundled with LS200® and MAGPIX® purchase. When using this software, experiment-specific settings and labels (sample IDs, standard curve concentrations, and controls) are entered into the system before sample acquisition begins. Standard curves and sample concentrations are calculated at the end of the acquisition step and are exported as a portable document file (.pdf).

Some MBAA reagent vendors sell their own multiplex assay analysis software. If using a second party software, such as Milliplex Analyst™ (Millipore®) or Bio-Plex Data Pro™ (Bio-Rad®, Hercules CA), raw data pertaining to bead counts and fluorescence intensity is imported as a comma-separated values (.csv) file, and experiment-specific settings (samples IDs, standard curve concentrations, and controls) are applied to the imported values. Standards and samples are then analyzed and results displayed in a table format, which can be exported to standard spreadsheet software (*see* **Note 16**). Figure 2 shows examples of standard curves produced by BeadView® software (v.1.03, Upstate) analysis software, and Fig. 3 is a set of sample data for quantifying a panel of six cytokines/chemokines in mouse sera.

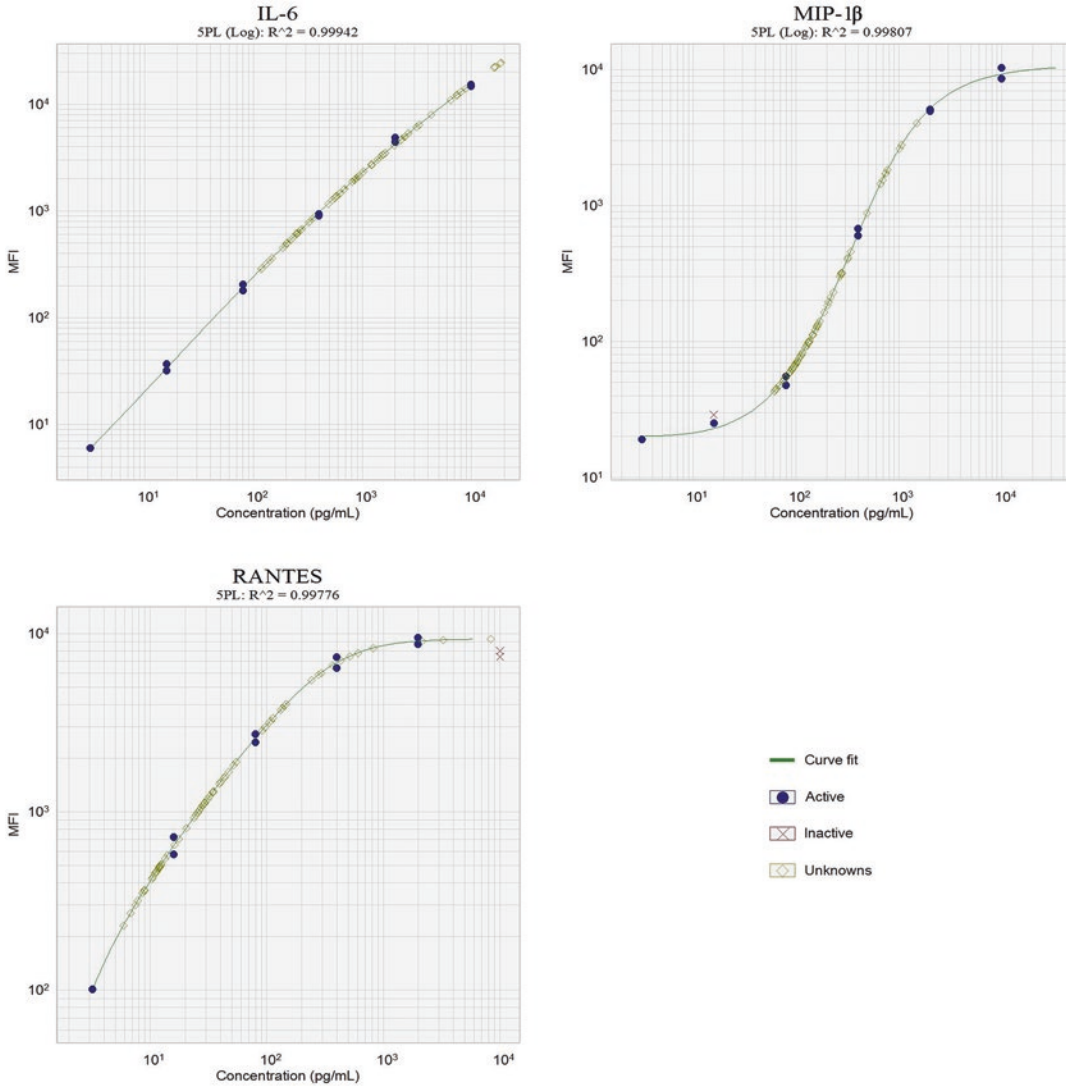


Fig. 2 Sample standard curves from a MBAA. Standards curves were generated using a Milliplex custom mouse cytokine/chemokine panel, and mouse serum samples. Grouped blue dots indicate duplicate standard dilutions, open diamonds indicate samples as extrapolated onto the standard curve, and red x's indicate standard curve dilutions which were omitted from curve calculation, due to low fluorescence intensity or variation between duplicates. Curves were generated using BeadView® software (v.1.03, Upstate)

4 Notes

1. The BD Cytometric Bead Array™ (CBA) is a similar technology available through BD Biosciences®. The basic principles of CBA are similar to Luminex® MBAA; however, instead of a single-use piece of equipment, CBA uses a standard dual laser flow cytometer (which can be equipped with accessories for automated sample acquisition and increased throughput) and

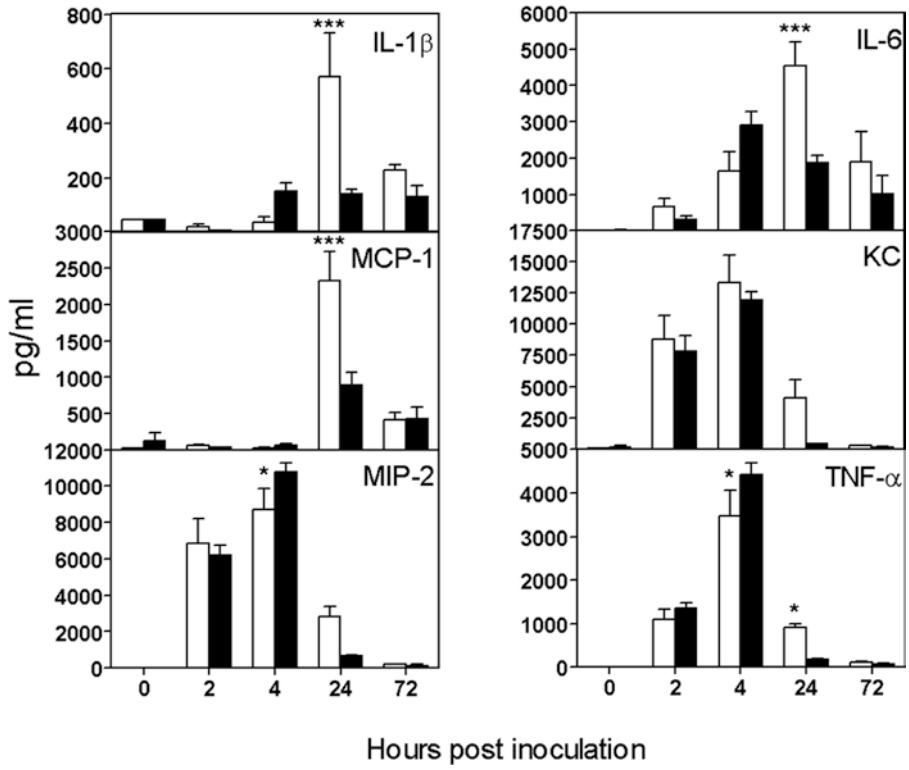


Fig. 3 Cytokine and chemokine levels in bronchoalveolar lavage fluid of A/J and C57BL/6 mice following i.n. inoculation with *A. baumannii.* Groups of A/J or C57BL/6 mice were i.n. inoculated with 2.5×10^7 CFU of *A. baumannii*. BAL samples were collected at 0, 2, 4, 24, and 72 h, and cytokine and chemokine levels were determined using the mouse panel of Fluorokine MAP Multiplex Kits (R & D Systems, Inc. Minneapolis, MN) on a Luminex® 100 IS instrument. Data are expressed as mean ± SEM of five mice at each time point. The detection limits of the assays were 2.5–15 pg/mL. $*P < 0.05$ and $***P < 0.005$ vs. C57BL/6 mice. From Qiu et al., Microbes and Infection 11 (2009) 946–955

specialized analysis software (FCAP Array 3.0™) that works in conjunction with BD FACSuite™ flow cytometer software. Custom kits are available for human, mouse, and rat, with a focus on cytokines/chemokines, growth factors, cell signaling, and isotyping. In addition, unconjugated CBA beads are available for the creation of custom bead/antibody complexes, similar to Luminex MAGPLEX™ beads.

Ultimately, Luminex®-compatible products are more widely used than BD CBA, and the associated partners and product lines cover a broader spectrum of both life sciences research areas and species range. This chapter will focus solely on Luminex®-compatible technology.

2. LS200® and MAGPIX® require different types of calibration/verification beads. These are purchased from Luminex®, and their batch-specific target values and expiration dates are imported into the analyzer's computer system and saved for

future use. Calibration/verification reagents typically have an expiration date of 18–24 months from the date of purchase.

3. To speed up this step, float reagents in a room temperature water bath for 10 min.
4. The ideal speed is one where maximal shaking occurs without splashing of well contents. It is recommended that the user determines the optimal speed for their shaker or rotator on a blank plate with 100 μL/well of water before beginning the assay. While a rotator is generally recommended in kit instructions, we use a plate rocker at 120 rpm without complications. Be sure to secure the plate with tape or elastic bands.
5. Omit this step if a premixed bead cocktail is supplied with the kit although custom-designed kits generally require the preparation of bead cocktail by the end user.
6. Polystyrene beads can stick together after long-term storage. While the analyzer is programmed to gate out any doublets (i.e., made up of two different bead types, which would give a false reading), sonication helps to minimize the total amount of doublets, thereby increasing the total bead count and the assay accuracy.
7. Alternatively, the user can calculate the exact number of beads needed for the assay and only mix those into the cocktail. This is recommended when only a small number of samples are being assayed. Undiluted beads, as well as the remainder of the kit contents, have a relatively stable shelf life at 4 °C, and some kits can be used months or years after the kit's expiration date.
8. Samples are diluted 1:2 in assay buffer, whereas standards and quality controls are diluted 1:2 in the buffer that was used to prepare the samples (i.e., serum matrix if serum samples are being assayed; bronchoalveolar lavage (BAL) fluid if BAL samples are being assayed; PBS or saline if lung homogenate supernatants are being assayed; complete tissue culture medium if tissue culture supernatants are being assayed).
9. We have assayed mouse cytokine and chemokine levels in the serum, BAL fluids, the supernatants of tissue homogenates (lung, spleen, liver, and skin), and tissue culture supernatants using the MBAA system. It is generally recommended to centrifuge the samples after thawing to remove precipitating materials before assay even if no precipitate is visible. The amount of precipitate varies from sample to sample, with larger amounts in lung homogenates, smaller amounts in spleen homogenates, and very small amounts in mouse serum, BAL fluid, and tissue culture supernatants. In addition, the tissue homogenate pellets tend to be loose, especially with lung and liver samples. Therefore, remove supernatant gently without dislodging the pellet.

10. An overnight incubation is recommended by the manufacturer for a higher assay sensitivity.
11. Always confirm that the plate has been locked onto the separator for 1 min before flicking the liquid off. This step of the process is most prone to catastrophic human error if beads are discarded into the sink with the wash buffer due to improper magnetic capture of the beads.
12. This incubation step is a convenient time to set up the analyzer, as it requires a 30 min warm-up time for the laser (LS200® only), as well as approximately 10–20 min for plate template setup.
13. If immediate acquisition is not an option, plates can be stored in the dark at 4 °C and read at a later time, after 30 min of shaking to resuspend the beads. Milliplex manufacturers recommend acquisition no more than 24 h later; however, we have read plates 72 h after incubation, with no complications or significant loss of fluorescence intensity.
14. The acquisition software features real-time display of events as they occur, including progress bars for acquisition of each bead type, as well as the bead classification gates. During setup, the bead ID numbers are assigned to the software, and predetermined gates are used to gate out any beads that don't fall within any expected regions. It is recommended that the user watches the acquisition of the first set of samples, to ensure that all bead ID numbers are properly entered. If an error was made during bead ID number assignment, then the improper gate would be displayed, and the corresponding bead populations would fall outside of the gate (*see* Fig. 4). If this occurs, the acquisition must be shut down immediately, and the assay settings reentered. Samples can be reacquired, even if the entire plate was acquired before the error was noticed. Unlike a flow cytometer, the LS200® and MAGPIX® will return the samples to their original wells after acquisition, albeit more diluted. Therefore simply use the handheld or automatic magnetic washer to remove the liquid, and replace with 150 µL PBS or sheath fluid. As a general rule, it is recommended that after acquiring a plate, the plate is sealed and stored in the dark at 4 °C until after initial data analysis is complete so that reacquisition remains an option.
15. Typically, 50–100 beads are acquired per bead type. However low bead counts are known to occur, and the system is set to "time out" if a certain period of time (typically 1 min) passes before the minimum number of beads is reached. In this case, the data may still be useable despite the low bead count. We have had instances where only 20 beads were acquired, and the ensuing results were in line with the next well containing a duplicate sample where 100 beads were acquired.

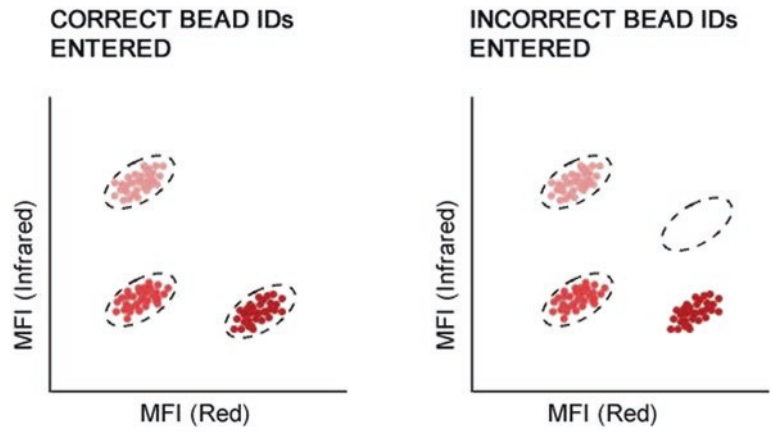


Fig. 4 Example of gate and bead locations from correctly and incorrectly entered bead ID numbers. Bead region gates (dotted lines) are pre-programmed into the software and are displayed based on the bead ID numbers entered by the user. Colored spots indicate all bead events and are displayed in real time on the screen during sample acquisition. The empty gate in the right hand panel indicates that the incorrect bead region was entered by the user. Note that this is a simplified portrayal of the data acquisition display screen

16. Before exporting data, the user should examine the standard curves generated by the software. The curves are typically sigmoidal, and each will have its own formula for a best fit, depending on binding behavior of each antibody type (best fit can be automatically assigned by the software or selected by the user). There are occasions where one or both of the standard curve duplicates fall outside of the calculated curve. *See* the RANTES standard curve in Fig. 2 as an example; in this instance, the top two points of the standard curve were removed, as their mean fluorescence intensity (MFI) registered lower than expected. The curve was recalculated using five points instead of six. If only one point needs to be removed (*see* MIP-1β standard curve, Fig. 2), the user should remove the point which results in a higher coefficient of determination (R^2) value when the curve is recalculated.

Samples should also be checked for their location on the standard curve. Ideally, all samples would fall within the linear range of the curve, but they can sometimes fall within the upper or lower plateau (*see* RANTES standard curve in Fig. 2). In the case of the upper plateau, a small difference in MFI between two duplicates can result in a very large difference in the calculated concentration. Therefore careful attention must be paid to the % CV between duplicates. Most analysis software will offer a display screen that shows the % CV of each set of replicates.

References

1. Khan SS, Smith MS, Reda D, Suffredini AF, McCoy JP Jr (2004) Multiplex bead array assays for detection of soluble cytokines: comparisons of sensitivity and quantitative values among kits from multiple manufacturers. Clin Cytometry 61B:35–39
2. Elshal MF, McCoy JP Jr (2006) Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA. Methods 38:317–323
3. Fulton RJ, McDade RL, Smith PL, Kienker LJ, Kettman JR Jr (1997) Advanced multiplexed analysis with the FlowMetrix™ system. Clin Chem 43:1749–1756
4. Keij JF, Steinkamp JA (1998) Flow cytometric characterization and classification of multiple dual-color fluorescent microspheres using fluorescence lifetime. Cytometry 33:318–323
5. Kettman J, Davies T, Chandler D, Oliver K, Fulton R (1998) Classification and properties of 64 multiplexed microsphere sets. Cytometry 33:234–243
6. Oliver KG, Kettman JR, Fulton RJ (1998) Multiplexed analysis of human cytokines by use of the FlowMetrix system. Clin Chem 44:2057–2060
7. Bellisario R, Colinas RJ, Pass KA (2000) Simultaneous measurement of thyroxine and thryotropin from newborn dried blood spot specimens using a multiplexed fluorescent microsphere immunoassay. Clin Chem 46:1422–1424
8. Taylor JD, Briley D, Nguyen Q, Long K, Iannone MA, Li M-S et al (2001) Flow cytometric platform for high-throughput single nucleotide polymorphism analysis. BioTechniques 30:661–669
9. de Jager W, te Velthuis H, Prakken BJ, Kuis W, Rijkers GT (2003) Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells. Clin Diagn Lab Immunol 10:133–139
10. de Jagger W, Prakken BJ, Bijlsma J, Kuis WJ, Rijikers GT (2005) Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. J Immunol Methods 300:124–135
11. DuPont NC, Wang KH, Wadhwa PD, Culhane JF, Nelson EL (2005) Validation and comparison of Luminex multiplex cytokine analysis kits with ELISA: determinations of a panel of nine cytokines in clinical sample culture supernatants. J Reprod Immunol 66:175–191
12. Codorean E, Nichita C, Albulescu L, Raducan E, Popescu ID, Lonita AC et al (2010) Correlation of xMAP and ELISA cytokine profiles; development and validation for immunotoxicoligical studies in vitro. Roum Arch Microbiol Immunol 69:13–19
13. Richens JL, Urbanowiscz RA, Metcalf R, Corne J, O'Shea P, Fairclough L (2010) Quantitative validation and comparison of multiplex cytokine kits. J Biomol Screen 15:562–1572
14. Searchable list of publications using MBAA (Luminex website). http://www.luminexcorp.com/publications
15. Qiu H, KuoLee R, Harris G, Chen W (2009) High susceptibility to respiratory *Acinetobacter baumannii* infection in A/J mice is associated with a delay in early pulmonary recruitment of neutrophils. Microbes Infect 11:946–955
16. Liu M, Xydakis A, Hoogeveen R, Jones P, Smith E, Nelson K et al (2005) Multiplexed analysis of biomarkers related to obesity and the metabolic syndrome in human plasma, using the Luminex-100 system. Clin Chem 51:1102–1109
17. Landlinger C, Preuner S, Willinger B, Haberpursch B, Racil Z, Mayer J et al (2009) Species-specific identification of a wide range of clinically relevant fungal pathogens by use of Luminex xMAP technology. J Clin Microbiol 47:1063–1073
18. Joslin E, Opresko L, Wells A, Wiley H, Lauffenburger D (2007) EGF-receptor-mediated mammary epithelial cell migration is driven by sustained ERK signaling from autocrine stimulation. J Cell Sci 120:3688–3699
19. Scholler N, Crawford M, Sato A, Dreschner C, O'Briant K, Kiviat N et al (2006) Bead-based ELISA for validation of ovarian cancer early detection markers. Clin Cancer Res 12:2117–2124
20. Lusk Pfefer TS, Timme R, Kase JA (2018) Identification of Brucella genus and eight Brucella species by Luminex bead-based suspension array. Food Microbiol 70:113–119
21. Kim K, Yang JS, Choi HB, Lee SH (2017) Detection of resistance to fluoroquinolones and injectable drugs among antituberculosis drugs by allele-specific primer extension on a microsphere-based platform. J Microbiol Methods 144:111–116
22. Juretschko S, Mahony J, Buller RS, Manji R, Dunbar S, Walker K, Rao A (2017) Multicenter clinical evaluation of the Luminex Aries flu

A/B & RSV assay for pediatric and adult respiratory tract specimens. J Clin Microbiol 55:2431–2438

23. Young S, Body B, Moore F, Dunbar S (2016) Multicenter evaluation of the Luminex ARIES HSV 1&2 assay for the detection of herpes simplex virus types 1 and 2 in cutaneous and mucocutaneous lesion specimens. Expert Rev Mol Diagn 16:1241–1249
24. Vignali DAA (2000) Multiplexed particle based flow cytometric assays. J Immunol Methods 243:243–255



Chapter 5

Preparation of the Low Molecular Weight Serum Proteome for Mass Spectrometry Analysis

Timothy D. Veenstra

Abstract

The ability to cure or manage many diseases is highly dependent on the ability to correctly diagnose them at the earliest possible stage. Diagnosis relies heavily on biomarkers whether these be visual symptoms or molecules found within samples acquired from the patient. For conditions that lack useful biomarkers, researchers are often faced with the task of sifting through very complex biological samples (i.e., serum, plasma, urine, tissue, cells, etc.) with the hope of discovering a small number of molecules that are exquisitely diagnostic for the condition of interest. One discovery strategy that has been frequently used is to fractionate the biological samples being studied into simpler aliquots that can be more easily characterized using existing technologies. One such fractionation method is to isolate a specific portion based on a specific property (i.e., size, phosphorylation state, charge, etc.) of the proteins within the sample. This method provides a simplified sample that can be characterized at a higher coverage level than the complex sample from which it was derived. This chapter details one of these methods, the extraction and analysis of the low molecular weight proteome of human serum.

Key words Low molecular weight proteome, Mass spectrometry, Serum, Protein depletion, Biomarkers

1 Introduction

The rapid rise in "omics" technologies has made it possible to characterize a significant portion of the molecular contents of cells in a reasonable time frame [1]. While the most significant advances have been made in genomics, technologies that afford wide proteomic coverage, particularly mass spectrometry (MS), have enable tens of thousands of proteins and their modified forms to be identified from single samples [2]. While having the capability of identifying thousands of proteins in a single sample is powerful, it can lead to situations in which a very large percentage of the acquired data clouds the important result sought in a study. For example, correctly identifying phosphorylated proteins in an entire cell extract can be challenging since most of the proteins

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_5, © Springer Science+Business Media, LLC, part of Springer Nature 2019

being analyzed will be non-phosphorylated. Not only does the confidence of correct identifications wain, but the instrumentation spends too much time analyzing proteins of little interest. While methods that analyze the entire proteome contents were the driving force behind improving proteomic technologies, the field is moving toward strategies in which samples are manipulated to target specific regions of interest within the proteome [3, 4]. These manipulations can include extracting proteins based on their function, modification, cellular location, charge, or size.

Identifying disease mechanisms is extremely challenging regardless of which omic (i.e., genomics, transcriptomics, proteomics, and metabolomics) approach is taken. The greatest challenge lies simply in the massive amount of data that is acquired using these technologies. Proteomics offers an additional challenge related to the dynamic flux of protein levels within samples acquired from specific (i.e., healthy) cohorts. Trying to determine if observed changes in protein levels is directly due to the disease or is simply normal cellular flux makes global proteomic studies very high risk. The reward, however, is very high if a useful drug-target or diagnostic/prognostic biomarker can be identified. Unfortunately, the recent era of global proteomics has resulted in too much failure and too little reward. Realizing this disparity has seen many investigators turn toward hybrid hypothesis/discovery-driven studies to reduce the complexity of biological samples prior to characterizing their proteomic content [5, 6].

One of the most important, yet challenging, sample types for proteomic studies is blood (either in the form of plasma or serum). It was recognized early on that a key step in analyzing blood samples for biomarker detection is to reduce the overall complexity of the sample [7–10]. Serum (as well as plasma) is an excellent candidate for complexity reduction since the most abundant proteins in serum comprise approximately 90–95% of the total protein content and are generally of moderate to high molecular weight [11–13]. In this study, a simple method to extract the low molecular weight portion of human serum, which eliminates a large portion of the highest abundant proteins, is described. While the term "low molecular weight" is relative, an arbitrary cutoff was established at 30 kDa; however, different molecular weight cutoff values can be used. Low molecular weight proteins in human serum have been linked with certain diseases such as dengue fever [14], diabetes [15], cancer [16], bladder ailments [17], and various cardiovascular ailments [18]. The procedure outlined in this chapter describes the removal of high-abundant proteins in serum (e.g., albumin, haptoglobin, and transferrin) resulting in the enrichment of low molecular weight and low-abundance proteins.

2 Materials

2.1 Preparation of Low Molecular Weight Serum

1. Standard human serum (NIST909c).
2. Protease inhibitor.
3. 50 mM ammonium bicarbonate (NH_4HCO_3), pH 8.4.
4. 25 mM ammonium bicarbonate (NH_4HCO_3)/acetonitrile (4:1 v/v).
5. Amicon Ultra-15 centrifugal filter units with a nominal molecular weight limit (NMWL) of 30 kDa and collection tube.
6. Falcon™ 15 mL Conical Centrifuge Tubes.
7. Centrifuge (e.g., Avanti J301, Beckman Coulter) with a fixed-angle (e.g., JA30.50) rotor.
8. Speed vacuum centrifuge.

2.2 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) Analysis to Confirm Removal of High Molecular Weight Serum Proteins

1. NuPAGE™ 4–12% Bis-Tris Gel.
2. Novex™ 10–20% gradient tricine gel.
3. Running buffer: 20× 3-(*N*-morpholino)propanesulfonic acid-sodium dodecyl sulfate (MOPS SDS).
4. XCell SureLock™ Electrophoresis Cell.
5. PowerEase™ 500 High-Voltage Power Supply.
6. Mark 12™ Unstained Molecular Weight (MW) Standard (Thermo Fisher Scientific).
7. 10× NuPage® Sample Reducing Agent (Thermo Fisher Scientific).
8. 4× NuPage® lithium dodecyl sulfate (LDS) sample buffer (Thermo Fisher Scientific).
9. 2.0 mL Safe-Lock polypropylene tubes.
10. 1.5 mL Safe-Lock polypropylene tubes.
11. Ammonium bicarbonate (NH_4HCO_3): 25 mM in water, pH 8.2.
12. SimplyBlue™ SafeStain (Thermo Fisher Scientific).

2.3 Enzymatic Digestion and De-salting

1. Porcine sequencing grade modified trypsin.
2. Incubator.
3. Speed vacuum centrifuge.
4. Bond-Elut C18 reversed-phase solid phase extraction (SPE) column (Agilent Technologies).
5. Ultrapure water (double distilled, deionized >18 Ω).
6. 2.0 mL Safe-Lock polypropylene tubes.
7. Dithiothreitol (DTT): 1 M in water.

8. Iodoacetamide: 1 M in water.
9. 50 mM NH_4HCO_3, pH 8.4.
10. Trifluoroacetic acid (TFA): 0.1% (v/v) in water.
11. Acetonitrile (CH_3CN): 60% (v/v) in water.
12. Methanol (CH_3OH): 60% (v/v) in water.

2.4 LC-MS/MS Analysis

1. Autosampler vials.
2. Branson sonicator.
3. Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer.
4. Capillary LC system (e.g., 1290 Infinity II c).
5. Fused-silica capillary tubing.
6. Jupiter C-18 column packing material.

3 Methods

3.1 Serum Preparation and Filtration

1. Slowly thaw a 2 mL aliquot of the standard human serum (NIST909c) by placing it on ice on the benchtop.
2. Add 20 µL of protease inhibitor to 200 µL of serum (*see* **Note 1**).
3. Dilute serum sample 1:1 by adding 220 µL of 50 mM NH_4HCO_3, pH 8.4.
4. Place an Amicon Ultra-15 filter in the filtrate retention vial.
5. Add 2–3 mL of 50 mM NH_4HCO_3, pH 8.4 to the filter.
6. Centrifuge the vial at 1000 × *g* for 2 min (*see* **Note 2**). Repeat this step.
7. Discard the filtrate and place the filter in a new filtrate retention vial.
8. Place 4 mL of 25 mM NH_4HCO_3/acetonitrile (4:1 v/v) in a Falcon™ 15 mL Conical Centrifuge Tubes.
9. Add 1 mL of the standard human serum (NIST909c) to the same Falcon™ 15 mL Conical Centrifuge Tubes from the above step. Mix gently by inverting five to ten times.
10. Transfer the diluted serum solution to the Amicon Ultra-15, and centrifuge at 3000 × *g* until approximately 90% of the solution has passed through the filter membrane.
11. Lyophilize the filtrate to complete dryness under vacuum.

3.2 Protein Analysis Using SDS-PAGE

1. Re-solubilize the lyophilized serum filtrate in 1 mL of 25 mM NH_4HCO_3.
2. Aliquot 20 µL of the solution to a 0.65 mL microcentrifuge tube.

3. Aliquot 7 µL of the NuPage LDS Sample Buffer into the tube.
4. Aliquot 3 µL of the NuPage Sample Reducing Agent into the tube.
5. Gently mix the sample and place it into boiling water for 10 min. Remove the sample from the boiling water bath, and allow it to cool to room temperature.
6. Remove the gel (either 4–12% Bis-Tris or 10–20% gradient tricine gel) from its package, and rinse it with water. Remove the protective tape at the bottom of the gel and the lane divider/holder from the top of the gel (*see* **Note 3**).
7. Prepare a MOPS solution by adding 50 mL of the concentrated (20×) stock solution to 950 mL of water, and mix gently.
8. Assemble the electrophoresis cell apparatus, as per the manufacturer's instructions, and add the diluted MOPS solution.
9. To one lane of the gel, add 10 µL of the Mark 12 Unstained MW Standard.
10. To another lane of the gel, add the serum mixture.
11. Cover the electrophoresis cell with the lid, attach the connectors to a power source, and run the gel at approximately 170 V for approximately 1 h (*see* **Note 4**).
12. When the dye front reaches the bottom of the gel, remove the gel, wash it in water, and stain with SimplyBlue SafeStain, as per the manufacturer's instructions, to confirm the removal of high molecular weight proteins from the serum sample.

3.3 Enzymatic Digestion and De-salting

1. Boil the remaining re-solubilized serum filtrate (i.e., approximately 1 mL) for 5 min, and let it cool to room temperature.
2. Add the 1 M DTT (prepared fresh) to the 1 mL serum filtrate for a final DTT concentration of 10 mM (*see* **Note 5**).
3. Heat the samples for 1 h at 56 °C and let it cool to room temperature.
4. Add the 1 M iodoacetamide (prepared fresh) to the serum filtrate for a final iodoacetamide concentration of 10 mM (*see* **Note 6**).
5. Incubate the sample for 1 h at room temperature.
6. Prepare a fresh solution of sequencing grade trypsin by solubilizing it in 50 mM NH_4HCO_3, pH 8.4 (*see* **Note 7**).
7. Add the trypsin solution to the serum filtrate so that the enzyme/protein ratio is 1:50 (*see* **Note 8**).
8. Digest the serum filtrate for 16 h at 37 °C with gentle shaking.

9. Acidify the serum filtrate with TFA to a final concentration of 0.1%.
10. Prepare an SPE cartridge by wetting the packing material with 2 mL of 60% CH_3OH/H_2O and flushing it with 2 mL of 0.1% (v/v) TFA.
11. Apply the digested serum sample to the cartridge, and remove the salts by flushing with 2 mL of 0.1% (v/v) TFA.
12. Elute the digested peptides with 400 μL of 60% acetonitrile/H_2O.
13. Lyophilize the samples at room temperature.

3.4 Sample Preparation for Mass Spectrometry Analysis

1. Add 100 μL of 0.1% (v/v) formic acid to the sample (*see* **Note 9**).
2. Sonicate each sample for 1 min, then gently vortex, and centrifuge briefly.
3. Transfer the re-solubilized samples to labeled autosampler vials and cap tightly.
4. Place on the LC autosampler for immediate analysis.
5. Peptide separation is performed using a 10-cm-long fused-silica capillary column (75 μm OD) prepared in-house with Jupiter C-18 (5 μm particle size, 300 Å) packing material (*see* **Note 10**). The column was coupled directly online with a mass spectrometer for protein identification (*see* **Note 11**).

4 Notes

1. The volume of serum used can vary depending on the source of the samples and scope of the study. Make sure to maintain the correct volume ratios for the addition of protease inhibitor and NH_4HCO_3.
2. The centrifugation times may vary based on filter pore size, temperature, and centrifugation speed. If the filtration unit has sodium azide added to the filter, flush it multiple times with water prior to use.
3. Either gel may be used for this step. Both gels gave near-identical results during initial tests.
4. Gel running times and voltages may be varied at this step based on the equipment used, manufacturers' recommendations, and gel types.
5. In lieu of using DTT, the protein disulfide bonds may be broken during the boiling step by using a commercially available solution such as Bond-Breaker tris(2-carboxyethyl)phosphine (TCEP) solution (Thermo Fisher Scientific, Waltham, MA).

This reagent is provided as a 0.5 M solution and provides a number of advantages over DTT or β-mercaptoethanol in that it can be used in either acidic or basic conditions and has no pungent odor compared to DTT and the breakage is almost irreversible, removing the need for alkylation.

6. Iodoacetamide is highly toxic and sensitive to light. Always use the proper precautions when using this reagent.
7. The trypsin solution was prepared by adding 1 mL of 50 mM NH_4HCO_3 (pH 8.4) to a vial containing 20 μg of trypsin. The vials are gently shaken and sonicated to solubilize the enzyme. Trypsin solution prepared in 50 mM NH_4HCO_3 may be frozen and stored at −20 °C but should be used within 2–3 weeks.
8. This protein/enzyme ratio may be varied based on the serum sample's protein concentration. The key is that the ratio should result in complete digestion of the proteins in the time allotted.
9. In the previous article on this technique published in 2013 [19], a pre-fractionation step using strong cation exchange (SCX) was utilized. While SCX can be used, the coverage afforded by recent advances in MS technology has dramatically increased the rate at which peptides can be identified. This increased rate generally obviates the need for pre-fractionation prior to LC-MS analysis.
10. Based on the investigator's preference and experience, they may wish to use their own preferred brand of column and packing material. The columns used for this study were prepared in-house with a length of 10 cm and used at a flow rate of 0.5 nL/min.
11. The type of mass spectrometer that can be used to collect the proteomic data varies based on the user's preference. It is recommended, however, that a mass spectrometer capable of collecting MS^2 and MS^3 data with high sensitivity, dynamic range, and fast duty cycle be utilized to maximize sample coverage.

Acknowledgments

This project has been funded in whole or in part with funds from Maranatha Baptist University. The mention of trade names, commercial products, or organizations does not imply endorsement by Maranatha Baptist University.

References

1. Karczewski KJ, Snyder MP (2018) Integrative omics for health and disease. Nat Rev Genet 19:299–310
2. Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. Nature 537:347–355
3. Matsumoto M, Nakayama KI (2018) The promise of targeted proteomics for quantitative network biology. Curr Opin Biotechnol 54:88–97
4. Paik YK, Omenn GS, Hancock WS et al (2017) Advances in the chromosome-centric human proteome project: looking to the future. Expert Rev Proteomics 14:1059–1071
5. Gu H, Ren JM, Jia X et al (2016) Quantitative profiling of post-translational modifications by immunoaffinity enrichment and LC-MS/MS in cancer serum without immunodepletion. Mol Cell Proteomics 15:692–702
6. Amit E, Obena R, Wang YT et al (2015) Integrating proteomics with electrochemistry for identifying kinase biomarkers. Chem Sci 6:4756–4766
7. Bouamrani A, Hu Y, Tasciotti E et al (2010) Mesoporous silica chips for selective enrichment and stabilization of low molecular weight proteome. Proteomics 10:496–505
8. Chertov O, Simpson JT, Biragyn A et al (2005) Enrichment of low-molecular-weight proteins from biofluids for biomarker discovery. Expert Rev Proteomics 2:139–145
9. Pisanu S, Biosa G, Carcangiu L et al (2018) Comparative evaluation of seven commercial products for human serum enrichment/depletion by shotgun proteomics. Talanta 185:213–220
10. Chen L, Zhai L, Li Y et al (2015) Development of gel-filter method for high enrichment of low-molecular weight proteins from serum. PLoS One 10:e0115862
11. Adkins JN, Varnum SM, Auberry KJ et al (2002) Toward a human blood serum proteome - analysis by multidimensional separation coupled with mass spectrometry. Mol Cell Proteomics 1:947–955
12. Dowling P, Ohlendieck K (2018) DIGE analysis of immunodepleted plasma. Methods Mol Biol 1664:245–257
13. Seong Y, Yoo YS, Akter H, Kang MJ (2017) Sample preparation for detection of low abundance proteins in human plasma using ultra-high performance liquid chromatography coupled with highly accurate mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 1060:272–280
14. Lim JM, Kim JH, Ryu MY et al (2018) An electrochemical peptide sensor for detection of dengue fever biomarker NS1. Anal Chim Acta 1026:109–116
15. Colombo M, Looker HC, Farran B et al (2018) Apolipoprotein CIII and N-terminal prohormone b-type natriuretic peptide as independent predictors for cardiovascular disease in type 2 diabetes. Atherosclerosis 274:182–190
16. Nagel M, Schulz J, Maderer A et al (2018) Cytokeratin-18 fragments predict treatment response and overall survival in gastric cancer in a randomized controlled trial. Tumour Biol 40:1010428318764007
17. Keay SK, Szekely Z, Conrads TP et al (2004) An antiproliferative factor from interstitial cystitis patients is a frizzled 8 protein-related sialoglycopeptide. Proc Natl Acad Sci U S A 101:11803–11808
18. Shimizu M, Doi S, Nakashima A et al (2018) N-terminal pro brain natriuretic peptide as a cardiac biomarker in Japanese hemodialysis patients. Int J Artif Organs 41:135–143
19. Waybright TJ, Chan KC, Veenstra TD, Xiao Z (2013) Preparation of the low molecular weight serum proteome for mass spectrometry analysis. Methods Mol Biol 1061:279–289



Chapter 6

Enriching for Low-Abundance Serum Proteins Using ProteoMiner™ and Protein-Level HPLC

Sophie Moggridge, Kelly M. Fulton, and Susan M. Twine

Abstract

The mammalian immune system acts to protect the body from harmful diseases ranging from cancer to infection. Differentially expressed proteins as a result of such an immune response can shed light on the mechanism of disease or serve as biomarkers. These biomarkers can be used in a diagnostic capacity or as correlates of protection following vaccination. Protein levels in the circulatory system are considered representative of the system as a whole, making serum an ideal matrix for surveilling immune responses. However, serum proteomics using mass spectrometry is extremely challenging due to the complexity of the matrix and the dynamic range of protein concentration. This chapter will describe two orthogonal enrichment strategies that can be used sequentially or in isolation to improve the identification of low-abundance serum proteins by mass spectrometry.

Key words Serum biomarkers, Biomarker, ProteoMiner™, Low-abundance protein enrichment, Dynamic range compression, Hexapeptide library

1 Introduction

Proteins exist in expansive networks that interact through complex signaling pathways. Expression is dynamic, and differential protein expression is often linked to disease states, such as cancer, mental illness, and infectious disease [1]. These differentially expressed proteins have the potential to be used as diagnostic determinants of health, known as biomarkers. Given the diverse utility of biomarkers, the continued development of methods for their identification is in the interest of regulatory bodies, researchers, and patients.

Serum is a preferred biological sample for biomarker identification and monitoring for two reasons. Firstly, as blood perfuses tissues, it gathers proteins from cellular responses in every organ system, making it a biological sample representative of overall health. Secondly, the collection of blood is minimally invasive compared to tissue biopsies and cerebral spinal fluid collection. Proteomic

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_6, © Springer Science+Business Media, LLC, part of Springer Nature 2019

approaches have great utility in the identification of serum biomarkers. However, the serum proteome is challenging to analyze due to its complexity, made up of an estimated 15,000 proteins [2], and the wide dynamic range of protein concentrations, which span more than ten orders of magnitude [3]. A small number of proteins, including albumin, immunoglobulins, transferrins, and haptoglobulins, account for more than 99% of the total protein content in serum [4]. These are typically of low immunological significance as biomarkers. The remaining 1% of the serum proteome represents hundreds of lower-abundance proteins, which have significant potential as disease biomarkers. Herein lies the main challenge in serum proteomics: identifying the lower-abundance proteins that may be more relevant for candidate biomarker discovery, across a large dynamic range of protein concentrations.

Proteomics research currently relies on mass spectrometry (MS) for protein identification. MS permits the analysis of the proteome in an unbiased manner because it has the capacity to identify any protein with a known mass and amino acid sequence. However, samples with a range of protein concentrations are challenging to analyze by MS because the signals from higher-abundance peptides can suppress the signal of lower-abundance peptides. There are various methods to improve serum protein identification by MS, typically involving depletion of the higher-abundance protein (e.g., immunoaffinity depletion) or simplification of the samples (e.g., fractionation). Immunoaffinity depletion uses antibodies to selectively remove targeted higher-abundance proteins, such as albumin. Despite the published success of this approach, it has some drawbacks, including limitations in antibody targets, and increasing evidence of nonspecific binding and sample loss due to co-elution of proteins that associate with higher-abundance proteins [5–7]. Fractionation separates a single complex sample into several less complex samples, in an attempt to improve protein identification. There are several ways to accomplish this, such as ion exchange chromatography (IEX), size exclusion chromatography (SEC), multidimensional liquid chromatography (MDLC), and polyacrylamide gel electrophoresis (PAGE). Fractionation tends to be lower throughput and requires significantly more analysis time.

Here we will describe two orthogonal approaches for serum simplification that can be used sequentially or in isolation to first deplete higher-abundance proteins in an unbiased manner and then fractionate to further reduce the complexity of the samples, prior to protein identification using LC-MS/MS. The first approach depletes higher-abundance serum proteins using a commercial bead-based peptide ligand library (ProteoMiner™, Bio-Rad) to non-specifically reduce the dynamic concentration range of the sample. The sample is mixed with hexapeptide-bound beads, and proteins are adsorbed by weak force or bonding interactions with the ligands. Higher-abundance proteins saturate the available

binding sites and the majority will flow through to waste. Lower-abundance proteins will not saturate their binding sites so the majority of these will be retained [8] and can subsequently be eluted (Fig. 1). Serum samples depleted of higher-abundance proteins can then optionally be subjected to an orthogonal separation method, such as protein high-performance liquid chromatography (HPLC), to increase the proteome coverage. While the ProteoMiner™ method has been described elsewhere [9, 10], a detailed description and methods for a workflow optimized for unbiased MS-based identification of lower-abundance serum proteins has not been described. In this chapter, we described the use of ProteoMiner and possible subsequent protein HPLC, combined with downstream traditional proteomics for identification of serum protein biomarkers (Fig. 2).

2 Materials

All solutions must be prepared using analytical grade chemicals and deionized water, unless specified otherwise. If working with human serum, follow necessary biosafety precautions. Dispose of all chemicals and biohazardous materials in appropriate waste containers.

2.1 Serum

1. At least 200 µL of clarified serum (*see* **Note 1**).

2.2 ProteoMiner

1. Centrifuge spin columns containing ligand-bound bead slurry.
2. Wash Buffer: 150 mM NaCl, 10 mM NaH_2PO_4, pH 7.4.
3. 2 mL capped centrifuge tubes.

2.2.1 ProteoMiner™ Protein Enrichment Kit, Small Capacity (Bio-Rad) (*See* **Note 2**)

2.2.2 ProteoMiner™ Sequential Elution Reagents (Bio-Rad)

1. Elution Reagent 1: 1 M sodium chloride, 20 mM HEPES, pH 7.4.
2. Elution Reagent 2: 200 mM glycine, pH 2.4.
3. Elution Reagent 3: 60% ethylene glycol in water.
4. Elution Reagent 4: 33.3% 2-propanol, 16.7% acetonitrile, and 0.1% trifluoroacetic acid.

2.3 Sample Cleanup

1. *OrgoSol*-Detergent-OUT™ micro kit (G-Biosciences), containing proprietary solutions UPPA-I, UPPA-II, DO wash, SEED solution, and OrgoSol Buffer.
2. Ammonium bicarbonate (ABC): 50 mM in water.

2.4 Protein HPLC

1. Binary LC pump, capable of capillary flow rates, equipped with a temperature-controlled column compartment, autosampler, diode array detector (DAD), and fraction collector with control software, for example, Agilent 1260 HPLC system with ChemStation software.

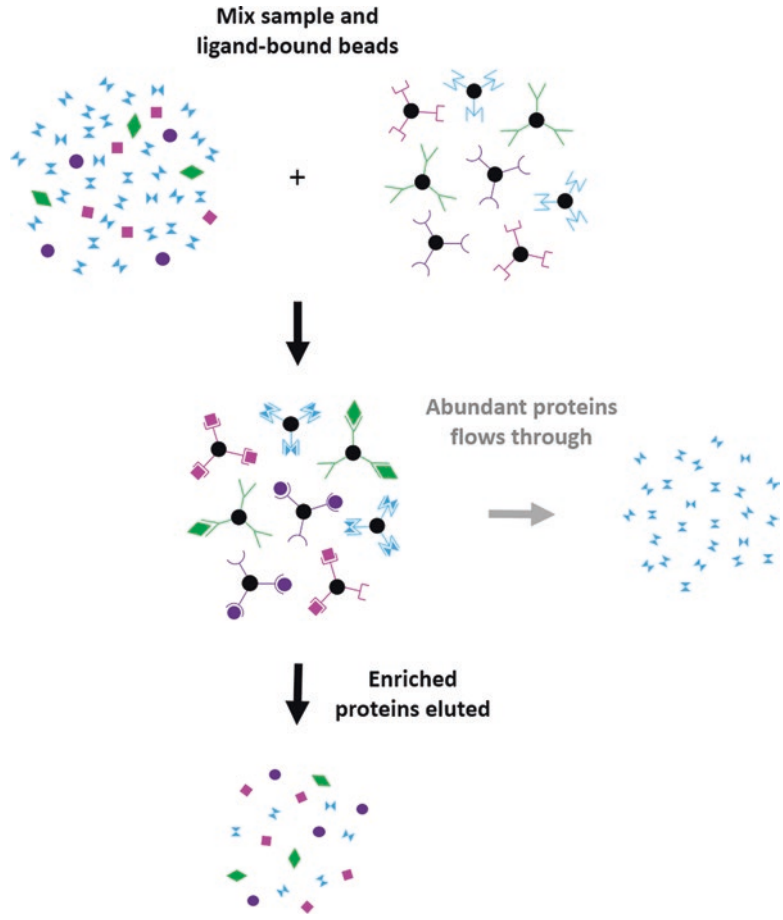


Fig. 1 Schematic showing the principle of serum depletion using a bead-based peptide ligand library. The sample is applied to the hexapeptide-bound beads, and protein is adsorbed. Higher-abundance proteins saturate binding sites and are not retained. Lower-abundance protein is retained and then eluted using elution buffers 1–4

2. 125 μm polyetheretherketone (PEEK) tubing (for lines 1–4, 6) (Fig. 2).
3. 75 μm fused silica tubing (for line 5 only).
4. *POROS* R2/10 2.1 × 100 mm HPLC column (Phenomenex).
5. Clear, fixed insert HPLC screw vials and caps with septum (*see* **Note 3**).
6. Mass spectrometer, such as a quadrupole time of flight (Q-TOF), with vendor-provided software.
7. HPLC solvent A: 0.1% formic acid in HPLC grade water, degassed.
8. HPLC solvent B: 0.1% formic acid in HPLC grade acetonitrile, degassed.

9. Dimethyl sulfoxide (DMSO).
10. Standard protein mix: 5×10^{-11} mol each of ribonuclease B (RNB), cyctochrome C (CC), myoglobin (MY), phosphorylase B (PHB), and 2.5×10^{-11} albumin (AL), enolase (EN). In a 15 mL disposable plastic tube, add 10 mL of water and 0.00619 g CC, 0.00750 g RNB, 0.00848 g MY, 0.1663 g AL, 0.02328 g EN, and 0.04860 g PHB (*see* **Note 4)**.

2.5 Protein Identification Using nLC-MS/MS

1. Reducing solution (10×): 40 mM DL-dithiothreitol (DTT) in 50 mM ammonium bicarbonate (*see* **Note 5**).
2. Alkylating solution (10×): 200 mM iodoacetamide (IODO) in 50 mM ammonium bicarbonate (*see* **Note 5**).
3. Sequencing grade modified trypsin (*see* **Note 6**).
4. Formic acid: 0.1% in water.
5. 96 Conical Well Nunc plate and heat sealing sheets or HPLC glass vials.
6. C8 µ-precolumn, 300 µm ID × 5 mm (e.g., Dionex).
7. HPLC BEH C18 column, 130 Å, 1.7 µm, 2.1 mm × 50 mm (e.g., Waters).
8. Nano-HPLC instrument, for example, NanoAcquity or M Class (Waters).
9. Mass spectrometer capable of MS/MS (e.g., LTQ-Orbitrap XL (Thermo Fisher Scientific) or Waters Synapt Q-TOF).
10. Search engine for the identification of peptides from MS/MS data. For example, MASCOT, X!Tandem, MyriMatch, Comet, and MS-GF+.

3 Methods

For the analysis of a single sample of serum, 200 µL of clarified serum is required. This can be processed by the ProteoMiner™ protein enrichment kit and protein HPLC fractionation sequentially before peptide identification by nLC-MS/MS. Peptide identification can also be carried out directly after ProteoMiner™ protein enrichment, without subsequent protein HPLC fractionation, if desired (Fig. 2). Unless otherwise stated, perform all steps at room temperature.

3.1 ProteoMiner Sequential Elution

1. Briefly centrifuge a spin column at 1000 × *g* for 30 s to recover all beads in the column.
2. Break off the bottom plug of a spin column, and place the column in a 2 mL un-capped collection tube (*see* **Notes 7–9**). Remove top cap of the spin column.

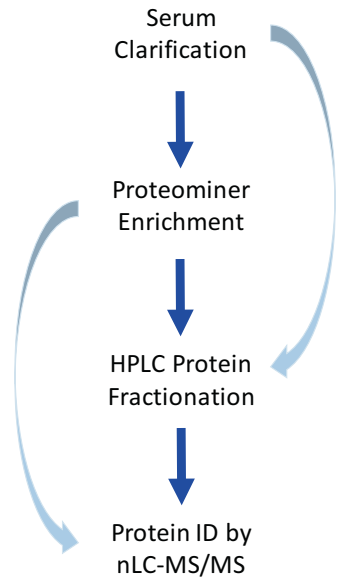


Fig. 2 Flowchart showing the workflow for protein depletion using bead-based peptide library and subsequent separation by HPLC. Peptide identification can be carried out directly after ProteoMiner™ protein enrichment or after subsequent protein HPLC fractionation

3. Centrifuge at 1000 × *g* for 30 s and discard the flow through.
4. Remove column from the collection tube, and replace the bottom plug of the spin column (*see* **Note 10**).
5. Pipette 200 μL of wash buffer into the column and replace the top cap (*see* **Note 10**).
6. Slowly rotate the column on a rotational shaker for 5 min.
7. Remove bottom plug and place the spin column in 2 mL tube. Remove the top cap of the spin column.
8. Centrifuge at 1000 × *g* for 30 s and discard the flow through. Replace the bottom plug of the spin column.
9. Repeat **steps 4**–7 once more using the same column. The column is now prepared to bind serum sample.
10. Pipette 200 μL of clarified serum into the spin column (*see* **Note 1**). Replace the top cap.
11. Rotate spin column on rotational shaker for 2 h.
12. Remove bottom plug and place the spin column in 2 mL tube. Remove the top cap of the spin column.
13. Centrifuge at 1000 × *g* for 30 s, and discard the flow through containing unbound proteins. Replace the bottom plug of the spin column.

14. Pipette 200 μL of wash buffer into the column and replace the top cap.
15. Slowly rotate the column on a rotational shaker for 5 min.
16. Remove bottom plug and place the spin column in 2 mL tube. Remove the top cap of the spin column.
17. Centrifuge at 1000 × *g* for 30 s and discard the flow through. Replace the bottom plug of the spin column.
18. Repeat **steps 13–16** twice more.
19. Pipette 200 μL of water into the column and replace the top cap.
20. Slowly rotate the column on a rotational shaker for 1 min.
21. Remove bottom plug and place the spin column in 2 mL tube. Remove the top cap of the spin column.
22. Centrifuge at 1000 × *g* for 30 s, and discard the flow through containing unbound proteins. Replace the bottom plug of the spin column.
23. Pipette 20 μL of Elution Reagent 1 into the column and replace the top cap.
24. Incubate column upright for 10 min, vortexing briefly at 2 min intervals.
25. Remove bottom plug and place the spin column in a clean 2 mL collection tube. Remove the top cap of the spin column.
26. Centrifuge at 1000 × *g* for 30 s and retain the eluate (Elution 1). Replace the bottom plug of the spin column.
27. Repeat **steps 22–25** twice more, collecting all three eluates in the same collection tube (Elution 1).
28. Perform **steps 22–26** using Elution Reagent 2, Elution Reagent 3, and Elution Reagent 4, sequentially.
29. The four elutions can either be stored at −20 °C or processed immediately as detailed in Subheading 3.2 (*see* **Note 11**).

3.2 Sample Cleanup

Sample cleanup must be performed on each of the eluted fractions in order to remove Elution Reagents 1–4, which are not compatible with downstream chromatography and mass spectrometry. Here we describe a cleanup procedure that uses the commercially available OrgoSol Detergent*OUT*™ kit.

1. Chill Orgosol Buffer at −20 °C for 1 h prior to use.
2. Transfer the entire 60 μL of each of the four elutions to a separate 1.5 mL capped Eppendorf tube (*see* **Note 12**).
3. Add 180 μL of UPPA-I solution to each elution. Vortex and then incubate sample on ice for 10 min.

4. Add 180 μL of UPPA-II solution to each eluate-UPPA-I sample. Vortex thoroughly.
5. Centrifuge all samples at 15,000 × *g* for 5 min (*see* **Notes 13** and **14**), and discard supernatant (*see* **Note 15**).
6. Spin the pellets at 15,000 × *g* for 30 s, and remove any residual supernatant using a pipette.
7. Add 40 μL of DO wash to the pellet without disturbing the pellet.
8. Centrifuge all samples at 15,000 × *g* for 5 min and discard the supernatant.
9. Add 25 μL of deionized water to each pellet (*see* **Note 16**).
10. Add 600 μL of OrgoSol Buffer and 3 μL of SEED to each tube. Vortex thoroughly.
11. Incubate the mixture upright for 30 min at −20 °C, vortexing for 30 s every 10 min.
12. Centrifuge samples at 15,000 × *g* for 5 min and discard the supernatant.
13. Air dry the pellet for 15 min in a laminar flow hood to evaporate any remaining liquid (*see* **Note 17**).
14. Store pellets at −20 °C until use (*see* **Note 18**).

3.3 Serum Protein Fractionation Using HPLC

1. The Agilent 1260 or similar HPLC capable of capillary flow rates should be set up as outlined in the schematic (Fig. 3). An Electrospray ionization (ESI) mass spectrometer can be used to monitor protein elution from HPLC column.
2. A *POROS* R2/10 2.1 × 100 mm should be connected as indicated in Fig. 3.
3. Degas HPLC solvents A and B (*see* **Note 19**).
4. Purge HPLC solvent A line on Agilent for 5 min at 2.5 mL/min.
5. Purge HPLC solvent B line on Agilent for 5 min at 2.5 mL/min.
6. Turn on the flow at 800 μL/min with 0% B for 10 min (*see* **Note 20**).
7. Set up an HPLC method using vendor software, to include equilibration of column, loading, and elution of protein samples. We suggest the following parameters: flow rate, 800 μL/min; injection volume, 10 μL; draw speed, 10 μL/min; eject speed, 10 μL/min; equilibration time, 1.2 s; sample flush out factor, 1.3× injection volume; column temperature, 55 °C; gradient time, 22 min (0% HPLC solvent B from 0 to 2 min, 0–1% HPLC solvent B from 2 to 3 min, 1–50% HPLC solvent B from 3 to 15.5 min, 50–85% HPLC solvent B from 15.5 to

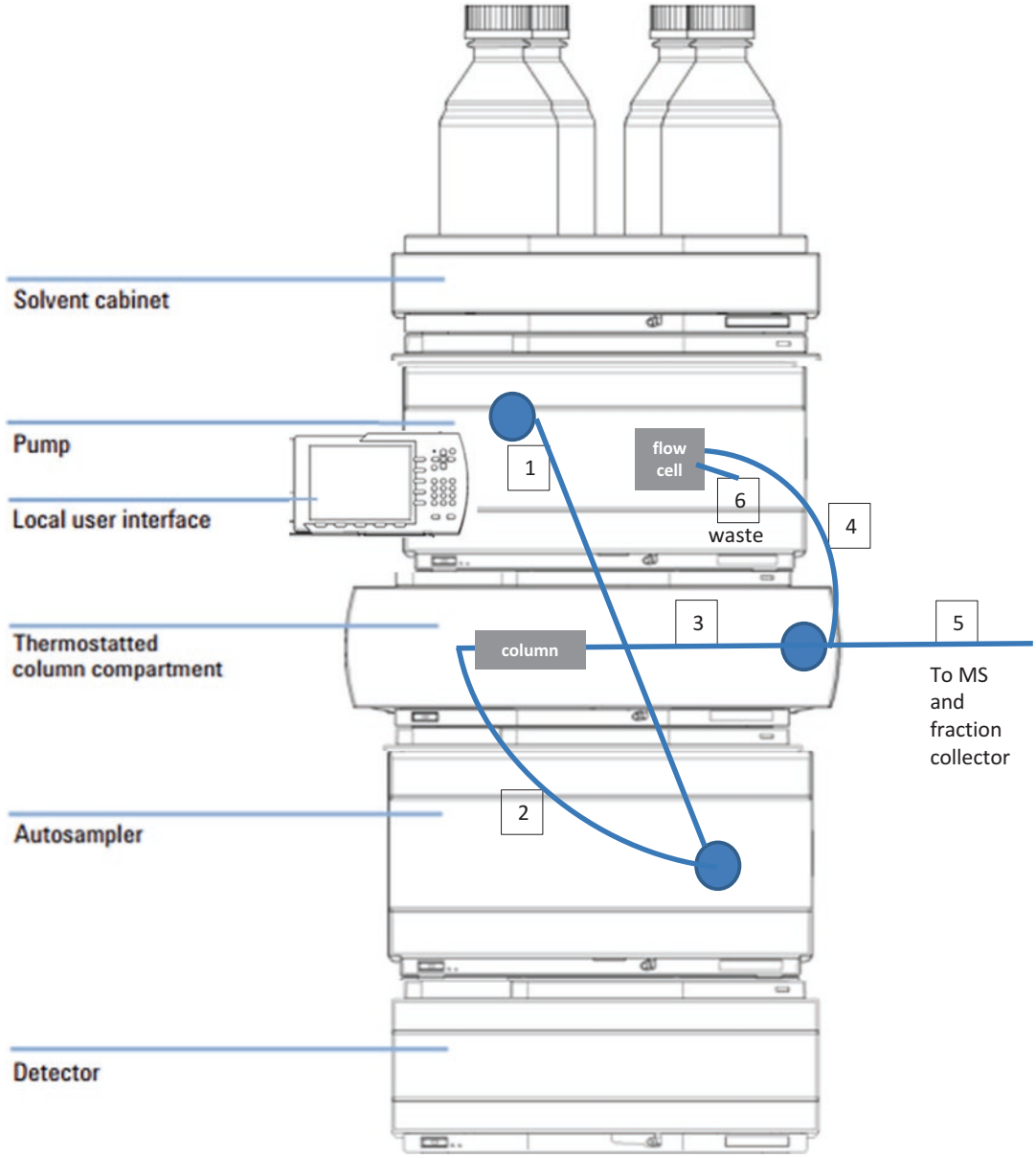


Fig. 3 Schematic indicating the setup of capillary flow HPLC for protein separation and monitoring by MS. Numbers relate to Subheading 2.4, which indicates the type of tubing recommended for various connections. Post column, line can be split to connect a fraction collector and MS for online monitoring

17 min, 85–0% HPLC solvent B from 17 to 22 min; max pressure 400 bar).

8. Set up a fraction collector to collect at 1 min intervals and connect as per Fig. 3 (*see* **Note 21**).
9. Allow flow to run for 5 min or until pressure has plateaued between 160 and 200 bar (*see* **Note 22**).

10. Elution of proteins can be monitored using an MS or by Ultraviolet (UV) monitoring.
11. Prepare sample list using mass spectrometer vendor software.
12. Inject 50 µL of standard protein mix to ensure system suitability. Figure 4 shows a typical chromatographic separation of standard protein mix (*see* **Notes 23** and **24**).
13. Resuspend Elution 1 pellet from Subheading 3.2, **step 14** in 10 µL of 5% DMSO, yielding an approximate concentration of 2 µg/µL.
14. Place 10 µL (~20 µg) of Elution 1 in a clean glass vial with cap. Place vial in HPLC autosampler.
15. Inject 10 µL of sample, and perform gradient as described in **step 7**.
16. Collect fractions in 1 min intervals between 7 and 17 min of the gradient (Fig. 5).
17. Repeat **steps 13–16** for Elution 2, Elution 3, and Elution 4 (*see* **Note 25**).
18. Evaporate all fractions to dryness using a vacuum centrifuge at 25 °C to remove organic solvents.

3.4 Protein Identification Using nLC-MS/MS

1. Resuspend pellets from Subheading 3.2 or fractions from Subheading 3.3 in 80 µL or 8 µL of 50 mM ABC, respectively.
2. Add reducing solution to 1× and incubate at 56 °C for 1 h.

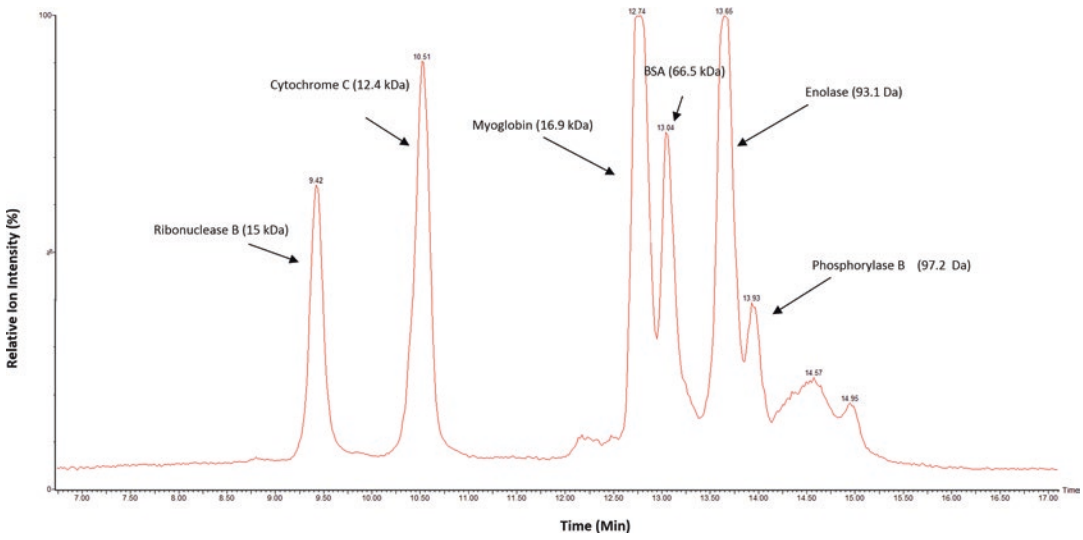


Fig. 4 LC-MS chromatogram showing separation of the standard protein mix. The names of the proteins are indicated. Elution times typically noted for comparison of reproducibility between injections

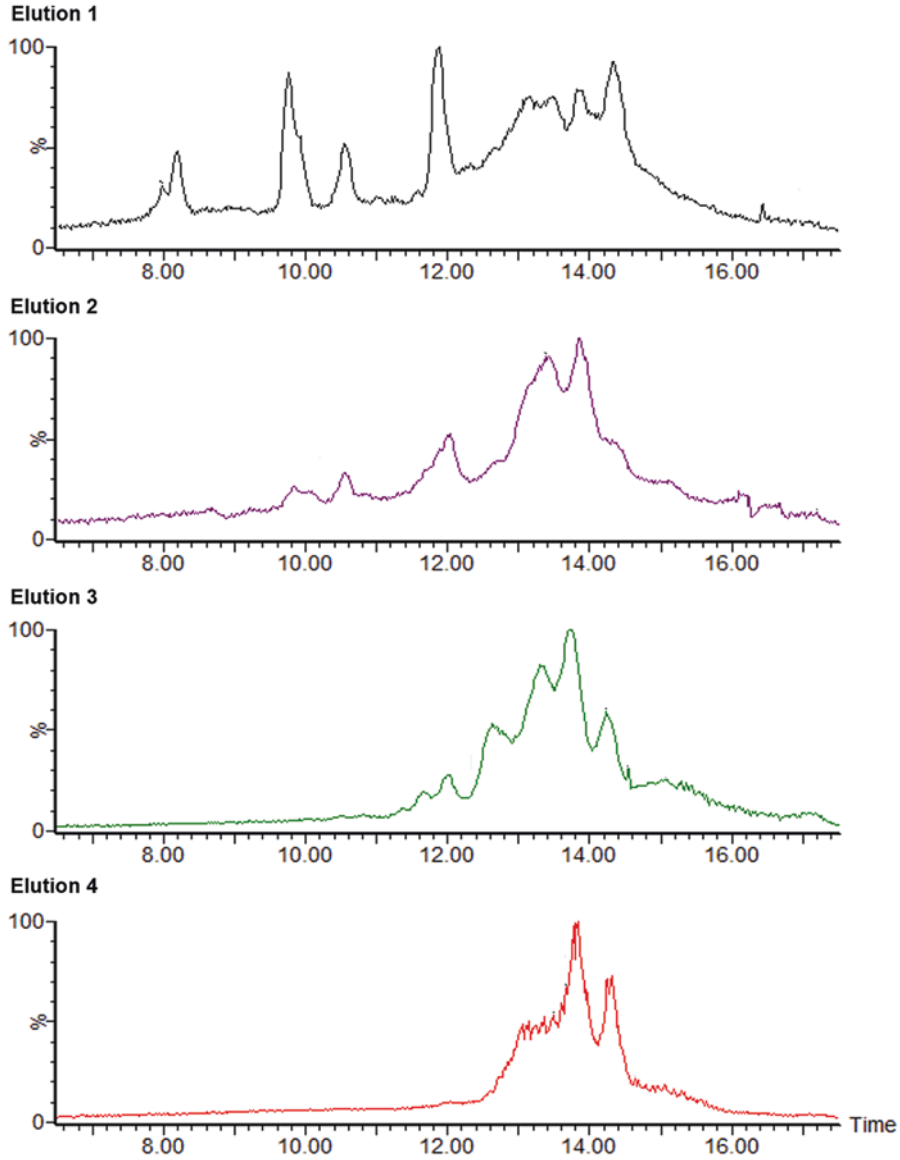


Fig. 5 LC-MS chromatograms showing protein HPLC separation of serum proteins depleted of higher-abundance proteins using ProteoMiner Sequential Elution Kit. Panels E1–E4 show the protein separation of fractions eluted with elution buffers 1–4

3. Recover condensation on tube lid with a 10 s spin in a bench-top mini centrifuge.
4. Add alkylating solution to 1×, and incubate at room temperature in the dark for 1 h.
5. Add trypsin at 30:1 protein to enzyme ratio (*see* **Notes 26** and **27**), and incubate at 37 °C for 18 h.

6. Recover condensation on tube lid with a 10 s spin in a benchtop centrifuge.
7. Add 0.1% formic acid to bring each elution to 100 μL or each fraction to 10 μL.
8. Centrifuge samples at 10,000 × *g* for 10 min to clarify.
9. Load 10 μL of each sample into a Nunc MS plate or clear glass vial with septum cap.
10. If using a plate, seal with a heat-sealable foil sheet. Centrifuge the plate for 10 s at 5000 × *g* to ensure samples are in the bottom of wells.
11. Set up an LC-MS system capable of tandem mass spectrometry (e.g., an LTQ-Orbitrap XL or Waters Synapt Q-TOF instrument) using data-dependent acquisition (DDA) in positive ion mode, with a scan range of 400–1600 m/z, and, for each scan, selecting top three parent ions for MS/MS.
12. Connect a nano-HPLC (e.g., Waters NanoAcquity or M Class), with C8 precolumn and C18 BEH column, to the mass spectrometer.
13. Inject 0.5 μg of protein at a flow rate of 0.4 μL/min.
14. Separate peptides using a nonlinear gradient of 1–45% HPLC Solvent B from 1 to 19 min, 45–85% HPLC Solvent B from 19 to 21 min, and 85–1% HPLC Solvent B from 21 to 22 min, and sustain at 1% HPLC Solvent B from 22 to 30 min for column re-equilibration.
15. Raw peptide and fragment data can be analyzed using one or several search engines, for example, Mascot, X!Tandem, Myrimatch, Comet, or MS-GF+ against the appropriate database. Typical database search criteria are as follows (*see* **Note 28**): Mascot ion score > 30, X!Tandem hyperScore > 40, Comet expect < 0.1, Myrimatch MVH > 40, and MS-GF+ SpecEValue < $1e^{-15}$.
16. Each protein is required to have, at minimum, two peptides identified by one search engine or one peptide identified by two search engines to confirm a positive identification. Redundancies were removed between fractions and replicates.

4 Notes

1. Serum samples must be clarified immediately prior to use in this protocol. To obtain 200 μL of clarified sample, spin 300 μL of serum in a 1.5 mL Eppendorf tube for 10 min at 1000 × *g*. The resulting sample will consist of a protein precipitate layer at the bottom of the tube, clarified serum in the middle, and lipids at the sample surface. Using a pipette, remove 200 μL of

the clarified serum sample taking care to avoid disturbing the precipitate or lipid layers. Place in a new Eppendorf tube, and keep on ice for up to 2 h prior to depletion steps. Discard the lipids and protein precipitate.

2. The lyophilized elution reagent (8 M urea, 2% CHAPS) and rehydration reagent (5% acetic acid) supplied with the ProteoMiner Small-Capacity Kit are *NOT* used in this protocol. The four sequential elution reagents are purchased separately from the kit.
3. To minimize sample loss by adsorption to the glass vials, it is recommended that the HPLC vials be siliconized. Surfacsil™ (Thermo Scientific) is a commercially available reagent that can be used to silanize vials, according to the manufacturer's instructions.
4. Standard protein mix includes five proteins of distinctly different masses and retention times. This mix is to be used as a system suitability test to ensure the HPLC is performing adequately prior to sample fractionation.
5. This solution must be prepared fresh and cannot be stored.
6. Trypsin must be mass spectrometry grade and preferably resistant to autolysis. Enzyme activity can be impacted by consecutive freeze-thaw cycles. Therefore it is recommended that the number of freeze-thaw cycles be limited to two.
7. Do not discard the top column cap and bottom column plug as they will be used throughout the protocol.
8. Once the plug is removed, liquid will flow through. Only remove the plug when the column is situated above a collection tube. This is particularly important during the elution steps.
9. For column preparation, serum sample binding and washing steps, use a single collection tube and discard flow-through material into appropriate waste container. For elution steps, use separate capped collection tubes for each of the four eluates.
10. When replacing the bottom plug on the column, ensure it is tight to prevent leakage and sample loss.
11. It may be desirable to retain a small aliquot of each elution to determine the protein concentration (e.g., by Bradford assay). However, in our hands using mouse serum as the starting material, each elution typically contains between 0.3 and 0.4 μg/μL of protein in each of the 60 μL elutions (~20 μg).
12. The conical collection tubes provided in the ProteoMiner Kit have a wide bottom. This makes it hard to see the protein pellet after precipitation steps. We recommend using 1.5 mL Eppendorf tubes, which make the pellet easier to see.
13. When precipitating proteins, sometimes the pellet may be translucent, small, or otherwise hard to see. To aid visualiza-

tion of the protein pellet, place the hinged tube into the microfuge with the hinge facing outward. This will ensure that the protein pellet always forms on the side of the tube below the hinge.

14. If the protein is not completely pelleted, additional centrifugation steps may have to be performed, to prevent disturbing the pellet when removing the supernatant. This will not negatively impact the protein pellet.
15. Do not attempt to pour supernatant off of pellet. Aspirate supernatant very slowly using a pipette to avoid sample loss.
16. The pellet may not fully dissolve in the buffer in this step but vortex to partially suspend the proteins in solution.
17. Be careful not to let the sample dry too much, as this will make resuspension of the pellet very difficult later on.
18. Depending on the degree of fractionation desired, proceed to Subheading 3.3 for protein-level HPLC fractionation prior to protein identification by nLC-MSM, or proceed directly to Subheading 3.4 for protein identification by nLC-MS/MS.
19. Degas with helium using a sparger or under vacuum.
20. If the detector becomes blocked, reversing the flow for 10 min will help dislodge blockages.
21. If using an offline fraction collector, program it to collect fractions at 1 min intervals.
22. Maintaining a constant, steady pressure in the HPLC system is critical for reproducible separation. It is normal for the pressure to increase slightly with increasing concentration of organic solvent flowing through the system.
23. The MS spray tip is very delicate, and care should be taken to avoid accidental damage. When attaching line 5 to the sprayer, be very careful not to crush the tip of tubing or the soft fitting.
24. HPLC and MS system performance may vary from day to day, but what is important is that the protein peaks are clearly visible above the baseline and that parameters such as peak width fall within a range that is acceptable for the system in use. We typically run a number of injections with standard proteins daily and note parameters such as HPLC elution time, peak width, etc. over time. This allows a record of system performance. The HPLC setup as outlined in this protocol typically sees distinct protein peaks, with peak width of between 0.2 and 0.3 min (CV of approximately 20%).
25. To maintain the column and prevent future blockages, flush column with 50% acetic acid after use. For short-term column storage (<1 week), flush the column with 100% Buffer A for

10 min. Column can be kept connected to the HPLC. For long-term column storage (>1 week), flush the column with 100% Buffer A (0.1% formic acid in H_2O) for 10 min before disconnecting the column. End plugs should be inserted and column stored at room temperature.

26. Lyophilized trypsin is resuspended in 50 mM Ammonium carbonate.
27. In our experience with mouse serum as the starting material, each elution contains approximately 20 μg of protein. Each of the ten fractions for each elution is assumed to have approximately 1/10 of the protein concentration, meaning 2 μg each.
28. We have typically used MASCOT for interrogation of peptide MS/MS data, with a peptide ion cutoff score of 30. For lower-abundance proteins, two peptides per protein are required to consider protein identification positive. With small, lower-abundance proteins, this can prove to be challenging. In some cases, we recommend manually inspecting the MS/MS data, to determine whether the predicted type *y* and *b* peptide fragment ions match those observed experimentally. In very large datasets, this is not practical; therefore it is essential to carry out false discovery rate analysis, in addition to designing experiments with appropriate numbers of technical and biological replicates. These factors should be considered together, to ensure resulting datasets of protein identifications can be analyzed with reasonable confidence.

References

1. Ahmed SS, Black S, Ulmer J (2012) New developments and concepts related to biomarker application to vaccines. Microb Biotechnol 5:233–240
2. Wiederin J, Ciborowski P (2016) Immunoaffinity depletion of highly abundant proteins for proteomic sample preparation. In: Ciborowski P, Silberring J (eds) Proteomic profiling and analytical chemistry, 2nd edn. Elsevier, Boston, pp 101–114
3. Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1:845–867
4. Tirumalai RS, Chan KC, Prieto DA et al (2003) Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics 2:1096–1103
5. Boschetti E, Righetti PG (2013) Low-abundance proteome discovery: state of the art and protocols. Newnes, Oxford
6. Di Girolamo F, Boschetti E, Chung MCM et al (2011) "Proteomineering" or not? The debate on biomarker discovery in sera continues. J Proteome 74:589–594
7. Yocum AK, Yu K, Oe T, Blair IA (2005) Effect of immunoaffinity depletion of human serum during proteomic investigations. J Proteome Res 4:1722–1731
8. Boschetti E, Righetti PG (2008) The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. J Proteome 71:255–264
9. Li L (2015) Dynamic range compression with ProteoMiner™: principles and examples. In: Posch A (ed) Proteomic profiling: methods and protocols. Springer, New York, NY, pp 99–107
10. Murphy S, Dowling P (2018) DIGE analysis of ProteoMiner™ fractionated serum/plasma samples. Methods Mol Biol 1664:109–114



Chapter 7

Identification of the Antigen Content of Electroimmunoprecipitates

N. Helena Beyer and Niels H. H. Heegaard

Abstract

Polyclonal antibodies including purified antibody fractions and animal or human antisera may react with unknown antigens or antigens other than their main specificity in reactions that are best visualized by gel electroimmunoprecipitation methods (e.g., when analyzing complex antigen mixtures). The great advantage of gel immunoprecipitation approaches is that each immunoprecipitate contains antigen in a pure form and that the precipitate is separated by position, shape, and size from other precipitates in the complex patterns of crossed immunoelectrophoresis. The identification of the antigen content of such immunoprecipitates is important but challenging because of the very stable, high-affinity complex formation leading to precipitation in the gels. Here, we present detailed step-by-step recipes for identifying the antigen content of electroimmunoprecipitates.

Key words Antigen identification, Mass spectrometry, Immunoelectrophoresis, Immunoprecipitates, Electroimmunoprecipitation, Antigen, Antibody, Dissolution of immunoprecipitates

1 Introduction: The Principle of Gel Electroimmunoprecipitation Techniques

Immunoelectrophoretic (IE) techniques are used for the quantitative determination of specific antigens that form immunoprecipitates with antibodies in gels [1]. The method works with most proteins if precipitating (polyclonal) antibodies are available and if the protein has a p*I* lower than that of the immunoglobulins. This ensures a differential electrophoretic mobility of antibody and antigen and hence electrophoretically mediated mixing in the gel. In comparison with the in-solution immunoprecipitation methods that are widely used in biochemistry, the IE techniques offer added possibilities for separation and quantification of multiple antigens in one operation. As an example, in crossed IE (CIE) [2], neither antibody nor sample

This article is dedicated to Niels H. H. Heegaard, who passed away suddenly during the preparation of this manuscript.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_7, © Springer Science+Business Media, LLC, part of Springer Nature 2019

needs to be pure. Thus, CIE can separate and quantify several antigens by using mixtures of antibodies in one analysis. For straightforward quantification of a protein in a crude mixture, so-called rocket IE (RIE) [3], a one-dimensional gel immunoelectrophoresis method, is simple, but requires a monospecific antibody.

In all electroimmunoprecipitation techniques, the identity of antigens in unknown precipitates may be an issue. An example is when more than one precipitate emerges with a supposedly monospecific antibody reacting with sample mixtures or when a particular precipitate appears in samples from a specific group of patients (e.g., a possible biomarker). To determine the identity of antigens in such immunoprecipitates, IE-immunoblotting methods have been devised [4–6], but these approaches are dependent on other antibodies and thus require an educated guess as to the identity of the antigen in question. Additional techniques for identifying unknown antigens in precipitates also are based on assumptions about the antigens and on the use of purified proteins and/or additional specific antibodies. These techniques include tandem-crossed IE [7], the use of intermediate gels with known antigens or specific antibodies [8], and immunofixation [9].

For identification of an unknown antigen, an orthogonal analytical method would be useful. An example is the excision of immunoprecipitates followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), which is an approach that gives an estimate of antigen molecular weight [10] and offers possibilities for further characterization. This approach does not appear to have been systematically optimized regarding the best method of extracting and recovering the antigen part of the excised immunoprecipitates. Also, electroimmunoprecipitates may be excised and analyzed directly by mass spectrometry after *in-gel* tryptic digestion. This is a simple and robust method, but due to the inaccessibility of some antigens in immunoprecipitates, it may suffer from insufficient tryptic fragmentation yields and thus a low sensitivity and antigen identification success rate.

Here, we present an optimized approach using RIE and CIE precipitates and SDS-PAGE with densitometry of protein bands and peptide mass fingerprinting (PMF) by mass spectrometry for identification of extracted proteins. The optimized extraction procedure is simple, sensitive, and robust. It allows the identification of antigens from electroimmunoprecipitates down to 135 ng of applied protein and may be used for characterization of antibody specificity.

2 Materials

Use MilliQ (MQ)-H_2O or similar for preparation of buffers and gels unless otherwise specified. Chemicals should be of high performance liquid chromatography (HPLC) grade or higher, except as noted.

2.1 Immuno-electrophoresis (IE)

2.1.1 Casting of Immune Gels

1. Leveling table.
2. Ethanol, 99% purity.
3. Tap water.
4. Detergent for dishwashing in the laboratory.
5. KimWipes or other lint-free paper towels.
6. Electric heating plate with magnetic stirring.
7. Water bath.
8. Steel blades or knives.
9. 5× Tris–tricine buffer (1 L) pH 8.6: 800 mL of MQ-H_2O, 26 mM tricine (*N*-Tris(hydroxymethyl)methyl glycine), 74 mM Tris-base (2-amino-2(hydroxymethyl)propane-1,3-diol), and 1.8 mM calcium lactate. Adjust pH with 5 N HCl to pH 8.6, and then bring final volume to 1 L with MQ-H_2O.
10. 1× electrophoresis buffer working dilution (2 L): 400 mL 5× stock Tris–tricine buffer, bring final volume to 2 L in MQ-H_2O.
11. 2.5 g agarose Human Serum Albumin (HSA) 100 Litex (BioWhittaker Molecular).
12. Glass plates (thickness of 1 mm) or Gelbond film (Lonza Cologne).

    Plate sizes: 5 × 5 cm, 5 × 7 cm, 7 × 7 cm, 7 × 10 cm, 10 × 10 cm, or 11 × 20.5 cm depending on the experiment (*see* **Note 1**).
13. Humidity chamber (plastic box with lid, line bottom with moist paper towels).
14. Gel puncher for sample application wells, 2.5 mm in diameter corresponds to 5 μL sample volumes.
15. Template for punching out sample application wells (*see* in **Note 2** and Fig. 1). For an additional example, please *see* Fig. 1.4 in [11].

2.1.2 Running of Immuno-Gels

1. Electrophoresis buffer from casting of gels.
2. Electrophoresis module, for details [11]:
   - Electrophoresis apparatus.
   - Cooling thermostat in circulating water bath.
   - Power supply with four channels, individually adjustable.
   - Test probe for direct measurement/adjustment of the potential gradient in the agarose gel.
   - Filter paper wicks, five pieces of Whatman No. 1 (Whatman) for low-voltage electrophoresis (2–3 V/cm, corresponding to a total of 70–110 V), must be the same width as immunoelectrophoresis plates for even distribution of voltage.

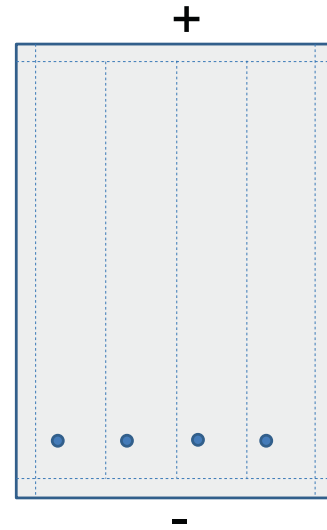


Fig. 1 Template for placement of wells and gel slabs of first dimension electrophoresis on 9 × 13 cm plate. The template is placed under the glass platesupporting the 1.5 mm thick gel. After first dimension electrophoresis the 5 mm gel along the edges is removed and the four gel slabs (2 × 11 cm) cut out individually. Placement of the anode for the first dimension electrophoresis is indicated

2.1.3 Washing, Pressing, and Drying of Immuno-Gels

1. Phosphate buffered saline (PBS) washing solution: 0.082 M NaCl, 0.043 M Na_2HPO_4, and 0.0098 M KH_2PO_4, in MQ-H_2O, pH 7.4.
2. Whatman No. 1, moistened (Whatman).
3. Soft absorbent tissue paper.
4. Glass plate (thickness of 8–10 mm).
5. Hot air stream for drying.

2.1.4 Visualization and Extraction of Electro-immunoprecipitates

1. Plate holder.
2. Container for staining solution.
3. Aqueous staining solution: 0.1% (w/v) Coomassie Brilliant Blue (CBB) R-250, in MQ-H_2O.
4. Standard staining solution: 0.1% (w/v) CBB R-250, 45% (v/v) methanol, and 5% (v/v) acetic acid, in MQ-H_2O.
5. Container for destaining solution.
6. Aqueous destaining solution: MQ-H_2O.
7. Destaining solution: 20% (v/v) ethanol (99% purity), 10% (v/v) acetic acid, in MQ-H_2O.

8. Soft absorbent tissue paper.
9. Extraction buffer: 0.06 M Tris–HCl and 10% (w/v) SDS in HPLC-grade H_2O, pH 7.

2.2 Running of SDS-PAGE

1. Spatula.
2. Power supply.
3. XCell Surelock Mini-Cell (Invitrogen Life Technologies).
4. 4–20% Tris-glycine gels (Invitrogen Life Technologies).
5. Low-range unstained SDS-PAGE molecular weight standard (Bio-Rad) or SeeBlue Plus2 Pre-Stained molecular weight standard (Invitrogen Life).
6. Sample buffer, final concentration in sample, total volume of 20 µL per sample: 8 mM Tris–HCl, pH 8.0, 1.6% (w/v) SDS, 2% (w/v) glycerol, 20 mM dithiothreitol (DTT), and 0.5 µg Pyronin G.
7. Tris-glycine gel running buffer: 25 mM Tris-base, pH 8.6, 192 mM glycine, and 0.1% (w/v) SDS.

2.3 Staining of SDS-PAGE

- CBB stain (*see* Subheading 2.1.4).
- CBB destaining (*see* Subheading 2.1.4).

2.3.1 CBB Staining

2.3.2 Mass Spectrometry-Compatible Silver Staining According to [12]

1. Fixing solution: 45% (v/v) methanol, 5% (v/v) acetic acid, in MQ-H_2O.
2. MQ-H_2O.
3. Sensitizing solution (50 mL): 0.02% (w/v) sodium thiosulfate ($Na_2S_2O_3$•$5H_2O$) MQ-H_2O.
4. Staining solution (50 mL): 0.1% (w/v) silver nitrate ($AgNO_3$), MQ-H_2O (prepared just before use, MQ-H_2O chilled at 4 °C).
5. Developing solution (50 mL): 2% (w/v) sodium carbonate (Na_2CO_3) and 15 µL of 37% (v/v) formaldehyde, in MQ-H_2O.
6. Stop solution (50 mL): 1% (v/v) acetic acid, in MQ-H_2O.

2.4 In-Gel Digestion with Trypsin

1. Lint-free wipes, such as KimWipes.
2. Ice.
3. Ethanol, 99% for wiping of surfaces and equipment.
4. Laminar air flow hood.
5. Bench-top centrifuge.
6. Vacuum centrifuge.
7. Incubator, preferably benchtop and shaking.

8. Microcentrifuge tubes: 1.7 or 0.65 mL SafeSeal, thin-walled PCR tubes (Sorenson BioScience).
9. Acetonitrile, HPLC-grade 100%.
10. H_2O: HPLC-grade.
11. Ammonium bicarbonate (NH_4HCO_3) buffer: HPLC-grade 100 mM NH_4HCO_3, in HPLC-grade H_2O.
12. Reducing solution: 10 mM DTT, 100 mM NH_4HCO_3, in MQ-H_2O.
13. Alkylating solution: 55 mM (10.2 mg/mL) iodoacetamide and 100 mM NH_4HCO_3, in MQ-H_2O.
14. Digestion buffer: 50 mM NH_4HCO_3.
15. Digestion solution with trypsin: 12.5 ng/μL trypsin, TPCK-treated, sequencing grade, and modified trypsin (Stratagene), in digestion buffer.
16. Resuspension solution (20 μL per sample).

Resuspension solution for matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS):

- 0.1% (v/v) trifluoroacetic acid in H_2O, HPLC-grade.

Resuspension solution for electrospray ionization mass spectrometry (ESI)-MS:

- 1% (v/v) HPLC or analytical-grade formic acid in HPLC-grade H_2O.

2.5 Sample Preparation for MALDI Mass Spectrometry

1. C_{18} Stage-tips (Proxeon Biosystems) or C_{18} Zip-Tips (Eppendorf AG) or POROS 20 R2 (Applied Biosystems).
2. GeLoader tips (Eppendorf AG).
3. Plastic syringe for micropurification.
4. KimWipes or similar lint-free wipes.
5. Washing buffer for MALDI-MS: 0.1% (v/v) trifluoroacetic acid in HPLC-grade H_2O.
6. Washing buffer for ESI-MS: 1% (v/v) HPLC or analytical-grade formic acid, in HPLC-grade H_2O.
7. Elution matrix for MALDI-MS: 6 g/L α-cyano-4-hydroxycinnamic acid, 33% (v/v) methanol, 33% (v/v) acetonitrile, 0.1% (v/v) trifluoroacetic acid (Rathburn, Walkerburn, UK), in HPLC-grade H_2O.
8. Elution for ESI-MS: 1% (v/v) HPLC or analytical-grade formic acid and 50% (v/v) methanol, in HPLC-grade H_2O.
9. Peptide standard I (Bruker) for calibration in MALDI-MS, applied according to the manufacturer's recommendation.
10. Polished steel target plate (Bruker).

2.6 Samples and Antibodies

1. Samples may be sera or other complex biological fluids.
2. Antibodies must be polyclonal for immunoprecipitation to occur.

3 Methods

For the entire workflow, refer to Fig. 2 for extraction of precipitates prior to *in-gel* digestion with trypsin (modified from [13]). Briefly, electroimmunoprecipitation is performed in agarose gels with subsequent CBB-based visualization of precipitates. Precipitates of interest are excised with a scalpel from either wet or dried gel and extracted overnight or preferably for 24 h (*see* **Note 3**) in SDS-containing extraction buffer. Contents of extracts are separated using SDS-PAGE. After visualization of proteins in the SDS-PAGE gel, protein bands are excised and *in-gel* digested with trypsin. After *in-gel* digestion, samples are micropurified and prepared for mass spectrometric analysis. Alternatively, the immunoprecipitates are directly *in-gel* digested and then subjected to mass spectrometry analysis (right track in Fig. 2). For an estimate of total time of analysis, *see* **Note 3**.

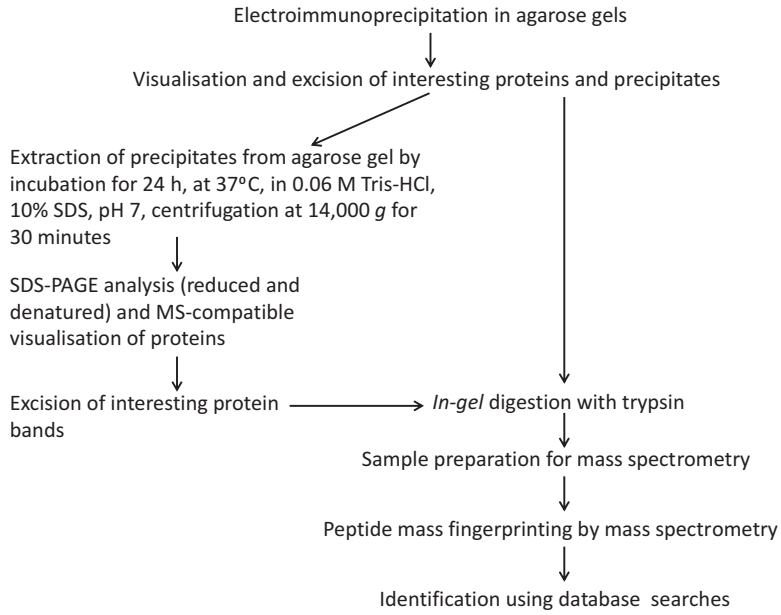


Fig. 2 The workflow involved in identification of electroimmunoprecipitated antigens. Visualized precipitates of interest are excised. Several precipitates may be pooled if the antigen is present in low amounts. Excision may be followed by extraction, SDS-PAGE analysis and *in-gel* digestion (left track of figure), or direct *in-gel* digestion with trypsin (right track of figure). After *in-gel* digestion, samples are prepared for peptide mass fingerprinting using mass spectrometry. Data are used for parent protein (antigen in original immunoprecipitate) identification by database searches

We found no signs to indicate when direct enzymatic digestion of electroimmunoprecipitates (right track) would be successful or when antigen identification could only be accomplished using extraction (left track). It may be as simple as being related to the size of the antigen relative to the size of the antibody, or it may be related to the nature and complexity of the precipitates (*see* **Note 4**). In most cases it will be necessary to apply an overnight extraction to get enough material for subsequent PMF-based antigen identification.

About 50% of the protein applied for immunoelectrophoresis was recovered in the corresponding band in the SDS-PAGE gel. The reason for the low recovery is thought to result from the nature of formation of the electroimmunoprecipitates making them exceedingly difficult to dissolve completely (*see* **Notes 4** and **5**).

The identification of antigens from electroimmunoprecipitates from CIE demanded more material compared to RIE antigen identification, which may result from overlapping of the precipitates. This could perhaps be overcome by the use of crossed line electroimmunoprecipitation using intermediate gels or other CIE modifications [1, 8].

3.1 Immunoelectrophoresis (IE)

3.1.1 Casting of 1% (w/v) Agarose Gels

1. Wash glass plates by hand for two cycles in detergent, and rinse with tap water.
2. Rinse in ethanol.
3. Wipe dry with KimWipes or similar lint-free wipes.
4. Place dry glass plates or Gelbond film (hydrophilic side facing upward) on the leveling table.
5. Add 2.5 g of agarose to 250 mL working dilution of electrophoresis buffer.
6. Mix by stirring on a heating plate until boiling and the agarose solution is homogeneous and clear.
7. Place in water bath at 56 °C. The solution is ready to use when the gel solution temperature is 56 °C. Gel solution stability is 3 months at 4 °C or a maximum of three cycles of boiling for use.
8. Add the proper volume of agarose gel (0.2 mL agarose per cm^2), ready to use (e.g. *max.* 56 °C) to a tube equilibrated to the temperature of the water bath (Subheading 2.1.1).
9. Add antibodies to the gel solution, e.g., 1/100 dilution (v/v, 2 μL/cm^2), depending on the antibody and sample.
10. Mix by gently turning tube upside-down 2–3 times.
11. Carefully pour agarose-antibody gel onto the plate.
12. The agarose gel must be evenly distributed.
13. Eliminate air bubbles by a gentle touch of (e.g., a pipette tip).
14. Leave for gelation.

15. After gelation, place in humidity box at 4 °C for a maximum of 24 h if not used for analysis right after gelation. Preferably, the gel should be used immediately after casting.

3.1.2 Sample Application and Running of Gel

1. Place the agarose gel on template for sample application wells.
2. Punch out sample application wells with sucking gel puncher.
   Punchers are used in diameters of 2.0, 2.5, 3.0, and 4.0 mm corresponding to sample volumes of 1–15 μL in gels of 1.5 mm thickness.
3. Dilute samples in electrophoresis buffer as appropriate for the punched hole in the gel.
4. Apply samples into the application wells by pipetting.
5. Add electrophoresis buffer to buffer chambers, 1 L per chamber.
6. Wet a total of ten filter paper wicks (Whatman No. 1 paper) cut into the same width as the immunoelectrophoresis plates (for even distribution of voltage) in electrophoresis buffer, five pieces in each buffer chamber.
7. Place gels on the electrophoresis apparatus.
8. Gently place the 2 × 5 filter paper wicks on 1 cm of the top and bottom of the gel, and ensure even contact between gel and filter paper wick.
9. For placing of gels and adjustment of voltage prior to immunoelectrophoresis:
   - Place electrophoresis lid and electrodes, and turn on voltage to adjust to the proper field strength (2–3 V/cm, corresponding to a total of 70–110 V). Ensure right polarity (movement usually toward the anode (+)).
   - Turn off power supply, and remove the electrophoresis lid and electrodes. Place a glass plate bridging the wicks and gel to ensure proper placing of gel and wicks and to prevent condensation of water on the gel. Reassemble the electrophoresis chamber by placing the lid and electrodes again.
10. Turn on voltage and perform immunoelectrophoresis.
    - For *RIE* according to (1): 2 V/cm for 16 h at 10 °C.
    - For *CIE* according to (2) performed in two dimensions:

First dimension: At 10 V/cm for about 1 h or until a bromophenol blue-labeled albumin marker has migrated about 4.5 cm from the application well.

Second dimension: After the running of the first dimension, cut gel, and place on the new glass plate, and allow the combined gel to gelate. The second dimension is performed at 2 V/cm for 16 h at 10 °C.

11. Stop immunoelectrophoresis by turning off the power supply, and remove gels.

3.1.3 Washing and Pressing of Agarose Gels (Pressing Times Are Not Critical)

1. Press gels for 20 min by squirting with MQ-H_2O and layering Whatman No. 1 paper (no air bubbles), soft absorbent tissue paper, and a heavy glass plate. Pressing increases removal of non-precipitated proteins and thus minimizes undesired background staining of the gel.
2. Wash for 20 min in PBS or isotonic NaCl.
3. Press gels as above.
4. Wash for 20 min in MQ-H_2O.
5. Press for 20 min as above.
6. Wash for 20 min in MQ-H_2O.
7. Press for 20 min as above.
8. Dry in a stream of hot air.

3.1.4 Staining and Destaining of Agarose Gels

1. Stain for 1 h using aqueous CBB stain.
2. Destain in water until the background gel without precipitates appears completely clear (*see* Fig. 3).

or

1. Stain for 15 min using standard CBB stain.
2. Destain in standard destaining solution until the background gel without precipitates appear completely clear (*see* Fig. 3).
3. Scan the gel in ethanol-wiped plastic wrap using a gel scanner or similar equipment.

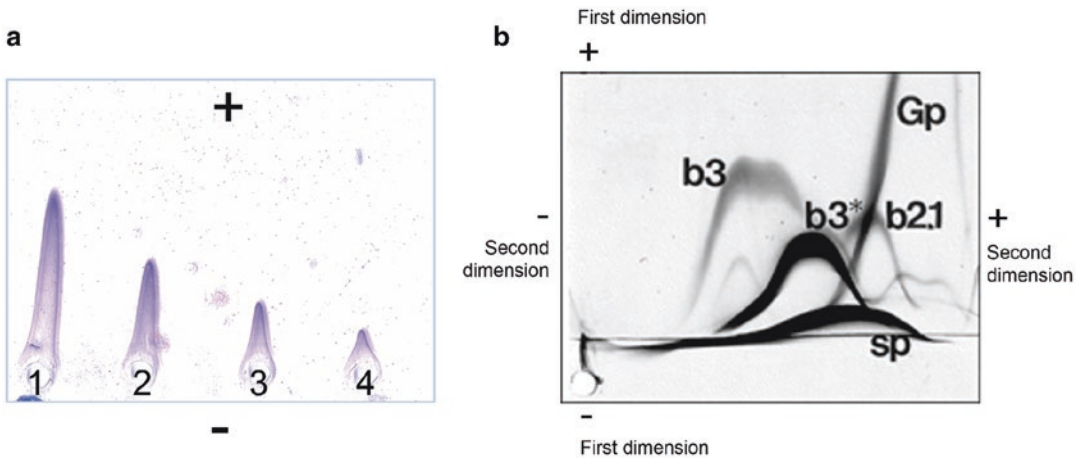


Fig. 3 Panel (**a**). Standard Coomassie Brilliant Blue staining of transthyretin/anti-transthyretin rocket immunoelectrophoretic precipitates. Antibody dilution was 1:150; applied serum was a dilution series, 1:20 (1), 1:40 (2), 1:80 (3), and 1:160 (4), according to [13]. Reproduced from [13] with permission from Elsevier. Panel (**b**). Crossed immunoelectrophoresis of 30 mg Triton X-100-solubilized human erythrocyte membrane proteins. The second-dimension gel contains 7 mL/cm^2 of a polyclonal rabbit anti-erytherocyte membrane protein antibody. *b2.1* band 2.1 protein (ankyrin), *b3/b3** band 3 protein (anion transporter, free and (asterisk) in complex with ankyrin), *Gp* glycophorin, *sp.* spectrins [14]

3.2 Extraction of Immunoprecipitates

1. Find IE gel, RIE or CIE, new or up to several years old (*see* **Note 6**).
2. Excise precipitates using a clean scalpel.
3. Place excised gel material in a microcentrifuge tube.
4. Incubate with 50–100 μL of extraction buffer at 37 °C for about 16–24 h (*see* **Note 7**).
5. Centrifuge for 30 min at 14,000 × *g*, according to [13].
6. Transfer supernatants to clean tubes.

3.3 SDS-PAGE

1. Make Tris-glycine running buffer.
2. Place 1 or 2 of the 4–20% Tris-glycine gels in the XCell Surelock Mini-Cell.
3. Add running buffer (600–800 mL).
4. Reduce 20 μL of the electroimmunoprecipitate extracts with 5 μL of reducing sample buffer.
5. Denature by heating at 100 °C for 5 min.
6. Load 15 μL samples in the gel wells.
7. Load molecular weight standards for SDS-PAGE if required.
8. Run electrophoresis at 120–150 V for 1–2 h according to [15].
9. Stop by turning off power supply.
10. Visualize protein bands by mass spectrometry-compatible staining, CBB staining, mass spectrometry-compatible silver staining [12], or other staining (*see* Fig. 4).
11. Scan the gel using a gel scanner or similar equipment.

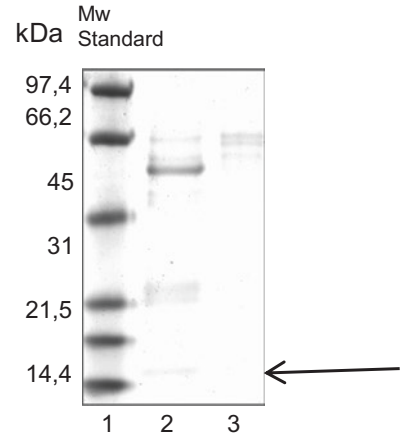


Fig. 4 Mass spectrometric compatible silver-stained SDS-PAGE of extracted proteins from immunoelectrophoretically produced precipitates from Fig. 2, Panel a. Lane 1, protein low molecular weight standard; lane 2, extract of precipitates; lane 3, blank from agarose gel. Arrow indicates the transthyretin protein band

3.4 In-Gel Digestion with Trypsin According to [12]

3.4.1 Washing

1. Excise protein bands or electroimmunoprecipitates of interest and place gel pieces in microcentrifuge tubes.
2. Add 200 µL HPLC-H_2O.
3. Mix for 10–15 min.
4. Remove liquid from gel pieces.
5. Add 50 µL of 100% acetonitrile.
6. Mix for 5–15 min (until gel plugs are white).
7. Remove liquid from gel pieces.
8. Dry down in a vacuum centrifuge (open lids) for a maximum of 30 min.
9. Turn on the heating block at 56 °C if reduction and alkylation is necessary.

3.4.2 Reduction and Alkylation

1. Add 50 µL of reducing solution to previously silver-stained protein bands of interest.
2. Place in heating block at 56 °C and leave for 45 min.
3. Cool to room temperature and remove liquid.
4. Add 50 µL of alkylating solution to the microcentrifuge tube. Leave in the dark for 30 min.
5. Remove liquid.
6. Add 10 µL of 100% acetonitrile (washing step).
7. Mix for 5–15 min (until gel plugs are white).
8. Remove liquid.
9. Add 15–25 µL of digestion buffer to the microcentrifuge tube.
10. Mix for 5 min.
11. Remove liquid.
12. Add 30 µL of 100% acetonitrile.
13. Mix for 5–15 min (until gel plugs are white).
14. Remove liquid.
15. Dry down in a vacuum centrifuge (open lids) for a maximum of 30 min.

3.4.3 Digestion

For bands excised from silver-stained or CBB-stained SDS-PAGE gels:

1. Add 20–30 µL of digestion solution with trypsin. Gel plugs must just be covered with liquid.

For electroimmunoprecipitates:

1. Add 1 µg trypsin per sample in digestion buffer. Gel plugs must just be covered with liquid.

For protein bands from SDS-PAGE or electroimmunoprecipitates:

1. Leave on ice for a minimum of 45 min but up to 2 h to ensure that trypsin is evenly distributed in the gel plug before digestion begins.
2. Remove excess liquid.
3. Add 20–30 μL or enough digestion buffer for the gel plugs to be covered.
4. Leave samples at 37 °C for 4–16 h to digest.
5. Pipette supernatant to a new microcentrifuge tube for further analysis.

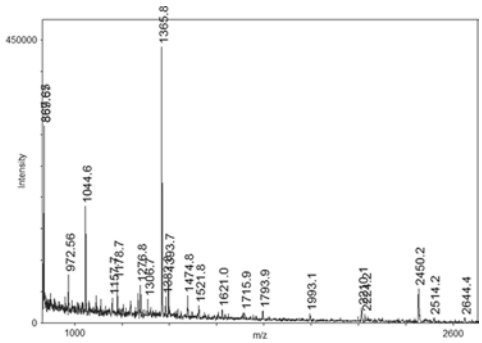
3.5 Sample Preparation for Mass Spectrometry

1. Micropurification of peptides is performed according to [16, 17]. Columns may be packed in the laboratory or be bought pre-made (Stage-tips (C18) or Zip-Tips (C18)).
2. Packing of columns in the laboratory may be done by restricting GeLoader tips and applying a suspension of POROS R20 matrix into the restricted GeLoader tips until a small column is formed [16, 17].
3. Activate and equilibrate the microcolumn, the Stage-tips (C18) or Zip-Tips (C18), according to the manufacturer's recommendations. Columns made in the laboratory may be activated with 50% methanol in 0.1% trifluoroacetic acid or 5% formic acid, followed by washing in 0.1% trifluoroacetic acid or 5% formic acid. They are then ready to use for micropurification of samples.
4. Micropurification may then be accomplished by application of the sample onto the microcolumn; wash with 10–30 μL of 0.1% trifluoroacetic acid or 5% formic acid.
5. After washing, sample peptides are eluted (see below).
6. For *MALDI-MS*: Elute samples with 0.8 μL of matrix (α-cynano-4-hydroxycinnamic acid, 6 g/L in 33% methanol, 33% acetonitrile, 0.1% trifluoroacetic acid) directly onto a polished steel target plate. Leave to air-dry and analyze.

   For *ESI-MS*: Elute samples with 1 μL of 50% methanol in 5% formic acid for ESI analysis. Samples may also be analyzed using online LC-ESI-MS on reversed-phase microcolumns, and thus micropurification may be omitted.

3.6 Mass Spectrometry

1. For reliable identification, use internal calibration and fragmentation.
2. Identification of peptides from PMF is based on online database searches in the latest available version of the NCBI database using the MASCOT search engine at http://www.matrixscience.com/cgi/nph-mascot.exe?1 (*see* Fig. 5).



gi|2098256
Mass: 13806 **Score: 74**
Expect: 0.0055
Queries matched: 6
Sequence Coverage: 74%

Chain B, Tertiary Structures Of
Three Amyloidogenic Transthyretin Variants And
Implications For Amyloid Fibril Formation

Matched peptides shown in **Bold Red**

1 **GPTGTGESKR PLMVKVLDAV RGSPAINVAV HVFRKAADDT WEPFASGK**TS

51 ESGELHGLTT EEEFVEGIYK VEIDTKSYWK **ALGISPFHEH AEVVFTANDS**

101 **GPRRYTIAAL LSPYSYSTTA VVTNPKE**

Start - End	Observed	Delta	Miss	Sequence
1 - 21	2210.1364	-0.0840	2	-.GPTGTGESKRPLMVKVLDAVR.G
22 - 34	1365.8295	0.0779	0	R.GSPAINVAVHVFR.K
35 - 48	1521.8027	0.0928	1	R.KAADDTWEPFASGK.T
36 - 48	1393.6995	0.0845	0	K.AADDTWEPFASGK.T
81 - 103	2450.2073	0.0095	0	K.ALGISPFHEHAEVVFTANDSGPR.R
104 - 127	2644.3640	-0.0108	2	R.RYTIAALLSPYSYSTTAVVTNPKE.-

Fig. 5 MALDI-MS, peptide mass fingerprint of *in-gel* digested TTR from Fig. 2. Identification of the antigen TTR from one excised band from SDS-PAGE, trypsin *in-gel* digestion, MALDI-MS, and subsequent database searches. Match indicating 74% sequence coverage

4 Notes

1. The size of the gel plate and consequently the immuno-gel chosen for an experiment depends on the number of samples and the complexity of both the polyclonal antibody and the samples. Generally, when doing RIE of a few samples and a specific antibody against a single antigen molecule (as the RIE in Fig. 3, Panel a), small- to medium-size gels (5 × 5–10 × 10 cm) are often sufficient for IE precipitates of the antigen. When more than four samples are to be applied to the gel, larger gels (10 × 10–10 × 21 cm) are recommended. When more complex samples/antibodies (e.g., serum samples and antiserum antibodies) or several dimensions are applied, medium gels may be used (5 × 7–10 × 10 cm).
2. Template for punching out sample application wells:

   Place the gelated agarose gel on the glass plate or Gelbond film on top of the template, and use the gel puncher to make sample application wells in the gel. Punchers typically used have diameters of 2.0, 2.5, 3.0, and 4.0 mm corresponding to sample volumes of 1–15 μL in gels of 1.5 mm thickness. The template may be applied to larger gels by replication.

3. Immunoelectrophoresis should run overnight for the best result. For extraction followed by SDS-PAGE: 1 day and overnight extration time is used, then 1 day for SDS-PAGE, staining, and excision of protein bands, 1 day for start of overnight *in-gel* digestion, and about 1 day for sample preparation for mass spectrometry and analysis, as this approach may result in many samples. Thus, including immunoelectromphoreses, a total of 5 days should be considered as analysis time from sample to result. Data analysis depends on the software available.
4. Immunoprecipitates created by electrophoretic mixing of antigens and polyclonal antibodies are very stable structures. This is due to the high-affinity and polyvalency of antibodies derived from hyperimmunized animals. Additionally, the immune complex stability is probably further enhanced by the enrichment of high-affinity interactions in an electrophoresis since antigen molecules binding with low affinity are removed by the electrical field when dissociating.

   This leaves very stably bound antigen covered with large antibody molecules binding to and cross-linking multiple epitopes in the final electroimmunoprecipitate. Such precipitates, especially in the dried gel, are stable indefinitely in the dark at room temperature. The accessibility of the antigen for further characterization is very limited and explains the relatively low yields upon direct trypsin-mediated fragmentation of electroimmunoprecipitates and the importance of optimizing immune complex dissociation when using the extraction procedure.
5. One factor is that the IgG bands on the SDS-PAGE gels were clearly stained much more intensely than the antigen bands, indicating that much more IgG than antigen was present in the extracts. This could reflect two things. If IgG and the antigen precipitate in a 1:1 molar ratio, then, in our examples, the stainability of the antibody part of the precipitate is ten times that of the antigen due to the difference in size.

   The much stronger intensity of the IgG protein bands on SDS-PAGE could also reflect that more than one IgG molecule is present per antigen molecule in electroimmunoprecipitates. This is likely with polyclonal antibodies that bind to multiple epitopes on antigen molecules.

   The success of the procedure probably varies with the strength of association between antibody and antigen. Thus, the sometimes difficult recovery is linked to the pronounced stability of the electroimmuno-complexes. It is well known that polyclonal antisera contain a spectrum of antibody affinities and electroimmunoprecipitation, as noted above, appears to select for the highest antibody affinities. Non-electrophoretic immunoprecipitation methods result in less stable immunoprecipitates as low-, medium-, and high-affinity antibodies will

all participate in the precipitation of the antigen and since incubation times are typically on the hour scale with no separation involved. In immunoelectrophoresis, however, precipitate is formed in overnight experiments and under the continued influence of an electrical field. Thus, it would be expected that immunoprecipitates from, for example, single radial immunodiffusion experiments [18], would be easier to dissociate than the corresponding antibody-antigen system precipitated by electroimmuno-techniques. Thus, the present procedure will also be readily usable for immunoprecipitates produced by other methods.

6. Both old and new electroimmunoprecipitates have been used (*see* Fig. 3). The method works even when gels have been stored for years, e.g., CIE of sera from systemic lupus erythematosus (SLE) patients from an earlier study [19] where serum amyloid protein (SAP) and transthyretin (TTR) were electroimmunoprecipitated with specific antibodies [13].
7. The optimal method was to excise the CBB-stained precipitate and extract it with a minimal volume of 0.06 M Tris–HCl (pH 7), 10% SDS, for 16–24 h at 37 °C. We found that at least 16 h of extraction is necessary. In addition, a temperature of 37 °C was better than higher temperatures, although extraction at room temperature could be investigated. Addition of ionic detergent, such as 10% SDS (w/v), was crucial for extraction, and less detergent resulted in a lower yield. Extraction solution could be either 0.06 M Tris–HCl at pH 7 or 0.1 M glycine, but Tris–HCl was preferred for compatibility with SDS-PAGE. The step of centrifugation was also crucial. Using this method, 135 ng TTR applied to an immunoelectrophoretic gel could be identified.

   To increase sensitivity, precipitates from up to eight separate CIEs may be pooled in a single microcentrifuge tube for extraction and subsequent PMF-based antigen identification.

References

1. Axelsen NH, Bock E (1983) Electroimmunoassay (rocket immunoelectrophoresis). In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 103–106
2. Grubb AO (1983) Crossed immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 113–124
3. Laurell CB (1966) Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. Anal Biochem 15:45–52
4. Bjerrum OJ, Hagen I (1983) Biomolecular characterization of membrane antigens. In: Bjerrum OJ (ed) Electroimmunochemical analysis of membrane proteins. Elsevier, Amsterdam, pp 112–126
5. Koch C, Skjodt K, Laursen I (1985) A simple immunoblotting method after separation of proteins in agarose gel. J Immunol Methods 84:271–278
6. Bjerrum OJ, Selmer JC, Lihme A (1987) Native immunoblotting – transfer of membrane-proteins in the presence of non-ionic detergent. Electrophoresis 8:388–397

7. Krøll J (1983) Tandem crossed immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 135–140
8. Axelsen NH (1983) Intermediate gel immunoelectrophoresis. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 141–149
9. Vandvik B, Nordal HJ, Vartdal F, Nilsen RE, Norrby E (1983) Imprint immunofixation of antibodies separated by agarose electrophoresis or by electrofocusing. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 33–37
10. Norrild B, Bjerrum OJ, Vestergaard BF (1977) Polypeptide analysis of individual immunoprecipitates from crossed immunoelectrophoresis. Anal Biochem 81:432–441
11. Svendsen PJ, Weeke B, Johansson B-G (1983) Chemicals, solutions, equipment and general procedures. In: Axelsen NH (ed) Handbook of immunoprecipitation-in-gel techniques. Blackwell, Oxford, pp 3–20
12. Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68:850–858
13. Beyer NH, Schou C, Houen G, Heegaard NHH (2008) Extraction and identification of electroimmunoprecipitated proteins from agarose gels. J Immunol Methods 330(1–2): 24–33
14. Heegaard NHH (1990) Immunochemical characterization of interactions between circulating autologous IgG and normal erythrocyte membrane proteins. Biochim Biophys Acta 1023:239–246
15. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685
16. Gobom J, Schuerenberg M, Mueller M, Theiss D, Lehrach H, Nordhoff E (2001) Alpha-cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics. Anal Chem 73(3):434–438
17. Kussmann M, Lässing U, Stürmer CA, Przybylski M, Roepstorff P (1997) Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of the neural cell adhesion protein neurolin purified by sodium dodecyl sulfate polyacrylamide gel electrophoresis or acidic precipitation. J Mass Spectrom 32(5):483–493
18. Mancini G, Carbonara AO, Heremans JF (1965) Immunochemical quantitation of antigens by single radial immunodiffusion. Immunochemistry 2:235
19. Sen JW, Recke C, Rahbek L, Skogstrand K, Heegaard NH (2002) Structural, quantitative and functional comparison of amyloid P component in sera from patients with systemic lupus erythematosus and healthy donors. Scand J Immunol 56:645



Chapter 8

Detection of Human Norovirus-Specific Antibodies Using the Luciferase Immunoprecipitation System (LIPS) Assay

Christine M. Tin and Stanislav V. Sosnovtsev

Abstract

The luciferase immunoprecipitation system (LIPS) assay is a liquid-phase immunoassay that quantitates antigen-specific serum antibodies by measuring luminescence emitted by the reporter enzyme *Renilla* luciferase (Ruc) fused to an antigen of interest. The LIPS assay can be utilized as a high-throughput and sensitive serological method for profiling serum antibodies recognizing diverse antigens. This chapter provides a detailed protocol for detecting human norovirus-specific serum antibodies with the LIPS assay.

Key words Human norovirus, LIPS assay, Serum, Antibodies, *Renilla* luciferase, Immunoassay

1 Introduction

Human noroviruses are the leading etiologic agents of acute gastroenteritis worldwide. In children under 5 years of age, noroviruses are estimated to cause 71,000 deaths annually, particularly in developing countries [1]. Human noroviruses are classified in the genus *Norovirus* of the family *Caliciviridae*. They are small, non-enveloped viruses with a ~7.7 kb positive-sense, single-strand RNA genome organized into three open reading frames (ORFs). Virus nonstructural proteins are encoded by ORF1, the major capsid protein VP1 is encoded by ORF2, and the minor capsid protein VP2 is encoded by ORF3 [2, 3]. The RNA genome is enclosed within an icosahedral protein shell that contains 180 copies of the VP1 protein. The VP1 protein is structurally subdivided into the shell (S) and protruding (P) domains, where the P-domain contains the most variable VP1 sequences exposed on the surface of the virus capsid [4–6].

The analysis of serum antibodies acquired during an adaptive immune response can provide insight into the etiology of the infection and protective immunity. Serological methods employed for characterization of norovirus infections include the enzyme-linked immunosorbent assay (ELISA) [7], hemagglutination and

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_8, © Springer Science+Business Media, LLC, part of Springer Nature 2019

hemagglutination inhibition assays [8, 9], western blotting [10], and immunoprecipitation [11]. However, assessing serum antibody responses to virus strains with antigenically diverse surface proteins can be difficult due to the genetic variability of many viruses, including human noroviruses [5, 6]. The luciferase immunoprecipitation system (LIPS) assay is a serological method recently developed to detect antigen-specific antibodies [12]. The LIPS allows high-throughput screening of antigen-specific serum antibodies in a liquid-phase format, which may be advantageous for serum antibodies to recognize surface epitopes on viral proteins that would otherwise be modified by interaction with a solid surface. The LIPS quantitates serum antibodies by measuring luminescence emitted by reporter protein *Renilla* luciferase (Ruc) fused to an antigen of interest. The fusion protein is expressed in mammalian cells, and crude lysates are harvested to extract the Ruc-antigen fusion protein. Antigen-specific antibodies recognize the Ruc-antigen, and protein A/G beads recognize the Fc region of the IgG antibody to capture antigen-antibody complexes. The LIPS assay has been successfully developed for profiling the antibody immune response against many viral proteins, as those of human norovirus, hepatitis C, HIV, herpes simplex virus, and Epstein-Barr virus [13–15]. This chapter presents the LIPS assay as a method for norovirus antigen-specific antibody detection and details a step-by-step method for its establishment.

2 Materials

2.1 General Supplies and Instruments for All Purposes (See Note 1*)*

1. Deionized water.
2. Molecular biology grade ultrapure water.
3. Ethanol alcohol, 99.9% purity.
4. Isopropanol.
5. 1.5 mL microcentrifuge tubes.
6. 15 and 50 mL conical centrifuge tubes.
7. 1× phosphate buffered saline (PBS) pH 7.4: 9.0 g/L NaCl, 144 mg/L KH_2PO_4, and 795 mg/L Na_2HPO_4 (anhydrous) in sterile-filtered water.
8. Refrigerated microcentrifuge.
9. Sorvall or similar centrifuge.
10. 50 mL high-speed centrifuge tubes.
11. 250 mL wide-mouth centrifuge bottles.
12. Disposable reagent reservoir.
13. Multichannel pipettor.
14. Handheld pipettor.

15. Pipette tips with filters.
16. Serological pipettes.
17. Class II laminar flow hood.

2.2 Plasmid Vector Cloning

1. pRen2 plasmid was a gift from Peter Burbelo (NIH, Bethesda, MD) and was the vector used for subcloning and expression [12].
2. High-fidelity PCR master mix.
3. Restriction endonucleases.
4. Deoxyribonucleic acid (DNA) ligation kit.
5. PCR strip tubes.
6. Tris–acetate–EDTA (TAE) stock solution (10×): 400 mM Tris–acetate and 10 mM EDTA in molecular biology grade water.
7. TAE buffer working solution (1×): dilute 10× TAE stock solution tenfold with deionized water.
8. 1% agarose gel.
9. Gel extraction kit.
10. Ethidium bromide.
11. DNA ladder.
12. Gel-loading dye.
13. Electrophoresis gel box.
14. 96-well thermal cycler.
15. Gel extraction tool or scalpel.

2.3 Bacteria Transformation, Growth Culture, and Plasmid Extraction

1. Chemically competent *Escherichia coli*.
2. SOC medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose.
3. 50 mg/mL kanamycin sulfate stock solution: dissolve 500 mg kanamycin sulfate in 10 mL ultrapure water.
4. Luria-Bertani (LB) agar plates with 50 μg/mL kanamycin: 1.0% tryptone (10.0 g/L), 0.5% yeast extract (5.0 g/L), 1.0% sodium chloride (NaCl; 10.0 g/L), 1.5% agar (15.0 g/L), and 50 μg/mL kanamycin.
5. 1 L disposable polycarbonate Erlenmeyer flask.
6. LB broth with 50 μg/mL kanamycin sulfate: 10.0 g/L NaCl, 10.0 g/L tryptone, and 5.00 g/L yeast extract in cell culture grade water. Add 1 mL of 50 mg/mL kanamycin sulfate stock solution to 1 L of LB broth.
7. Plasmid miniprep kit.

8. Plasmid maxiprep kit.
9. Disposable L-shaped cell spreaders.
10. Disposable inoculating loops and needles.
11. Shaking incubator for bacterial culture.

2.4 Mammalian Cell Culture and Transfection

1. COS1 cells (ATCC, Manassas, VA, USA).
2. Dulbecco's modified Eagle's medium (DMEM), high glucose, and GlutaMAX™ Supplement.
3. 10,000 U/mL penicillin/streptomycin stock solution.
4. Heat-inactivated fetal bovine serum (FBS).
5. COS1 complete growth medium: high glucose, GlutaMAX™-supplemented DMEM with 10% heat-inactivated FBS and 1% (100 U/mL) of 10,000 U/mL penicillin/streptomycin stock solution.
6. 0.05% Trypsin-EDTA.
7. Flat bottom cell culture 6-well plates.
8. Tissue culture treated T-150 flask.
9. Lipofectamine™ LTX with Plus reagent.
10. Opti-Mem™ reduced serum medium.

2.5 Extraction of Ruc-Antigen from Transfected Mammalian Cells

1. Ultrapure glycerol.
2. 1 M tris(hydroxymethyl)aminomethane(Tris)-HCl, pH 7.5.
3. 5 M sodium Chloride (NaCl).
4. 2 M magnesium Chloride ($MgCl_2$).
5. Triton X-100.
6. Halt protease inhibitor cocktail.
7. Disposable sterile bottle-top filter.
8. Cold lysis buffer A: 50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 1% Triton X-100, 50% ultrapure glycerol, and 1× Halt protease inhibitor in ultrapure water filtered with disposable sterile bottle-top filter using aseptic technique.
9. Cell scrapers.
10. Sonicator.

2.6 Serum Samples

For the data presented in this protocol, sera were collected from guinea pigs or from three conventionally raised Gottingen miniature pigs (minipigs). Guinea pigs and minipigs were immunized with human norovirus virus-like particles or mock-immunized under an animal protocol approved by the National Institute of Allergy and Infectious Diseases (NIAID) Division of Intramural Research Animal Care and Use Committee.

2.7 LIPS Assay Buffers, Reagents, and Instruments

1. Orbital shaker.
2. 96-well vinyl microtiter U-bottom plate.
3. 96-deep-well polypropylene microtiter plate.
4. Adhesive plate seals.
5. Buffer A: 50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, and 1% Triton X-100 in ultrapure water.
6. Pierce™ protein A/G Ultralink™ resin (Thermo Fisher Scientific, Waltham, MA, USA).
7. 96-well, white polystyrene, and opaque bottom microtiter plate.
8. Opaque 96-well filter HTS plate.
9. *Renilla* luciferase assay system.
10. Vacuum manifold.
11. Suction canisters.
12. Vacuum guard filter.
13. Microplate luminometer.

3 Methods

3.1 Construction of Expression Plasmid

A general schematic of the cloning strategy is outlined in Fig. 1, using BamHI and NotI as example restriction sites. The recombinant plasmid is constructed using restriction enzyme digestion and ligation of the corresponding amplified DNA fragment. The resulting construction is designed to express the antigen-coding sequence fused in the frame immediately downstream of the C-terminus of the Ruc reporter protein.

3.1.1 Preparation of Vector

1. The pRen2 vector (5.5 kbp) has a polylinker comprised of six unique restriction sites located directly downstream of the Ruc gene (Fig. 1). Digest 1 μg of pRen2 vector with two unique restriction enzymes that do not have sites in your gene of interest (*see* **Note 2**).
2. Run 1 μg of the digested plasmid in a nucleic acid gel designed for DNA electrophoresis (*see* **Note 3**).
3. Using a gel extraction tool or scalpel, excise the DNA band, and place the gel plug into a microcentrifuge tube.
4. Extract the linearized plasmid DNA using the gel extraction kit according to the manufacturer's instructions (*see* **Note 4**).
5. To determine purity of the DNA, run 10 μL of the gel-purified digest in a nucleic acid gel designed for DNA electrophoresis. A bright, clean band should be visible (*see* **Note 3**).

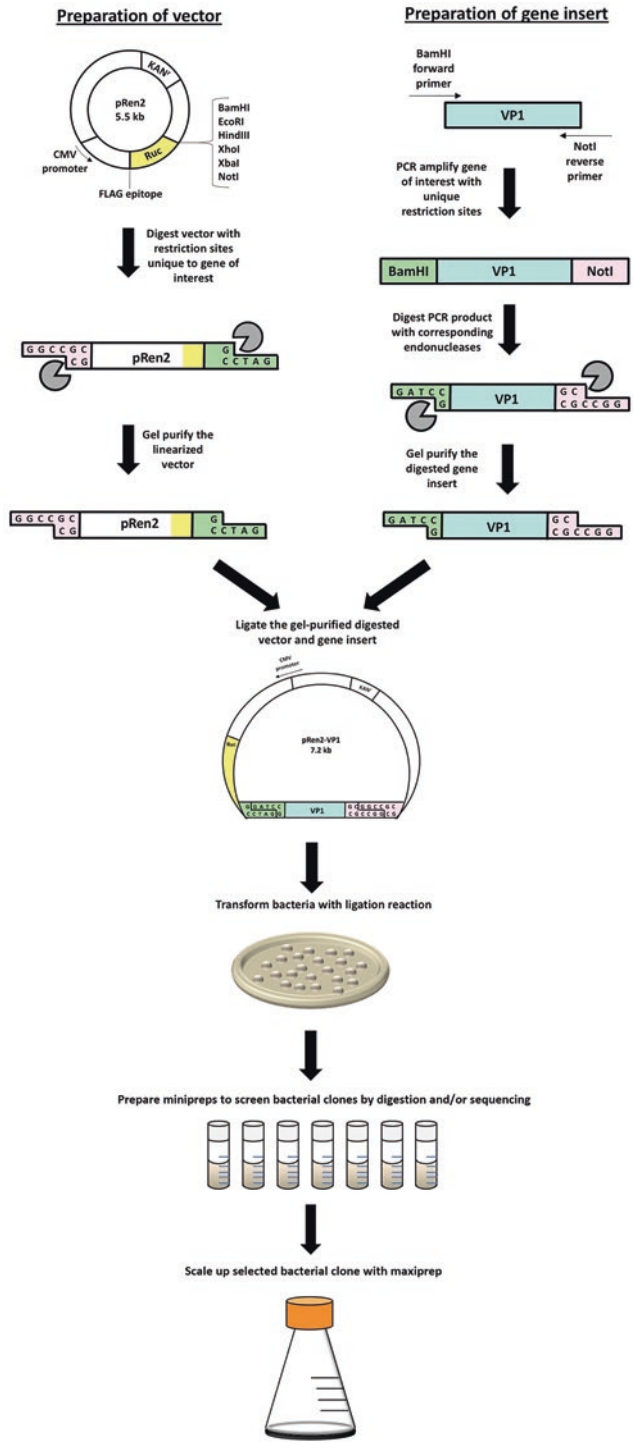


Fig. 1 Schematic of engineering the expression plasmid. The pRen2 vector is linearized by digestion with restriction enzymes generating two incompatible sticky ends in the plasmid's multiple-cloning site located downstream of the Ruc gene. The two recognition sites chosen in the plasmid vector for digestion should not be present in the gene of interest. The linearized vector is then gel-purified and stored at −20 °C until used.

3.1.2 Preparation of Gene Insert

1. Design appropriate oligonucleotide primers to introduce unique restriction sites to the 5′ and 3′ end of your gene of interest (*see* **Note 5**).
   (a) PCR amplify your gene of interest using 0.2 μM forward primer, 0.2 μM reverse primer, and PCR master mix in a 50 μL reaction. The following conditions for the thermal cycler are used: 94 °C for 2 min (1 cycle), 94 °C for 5 s, 55 °C for 30 s, 68 °C for 1 min/kbp (30 cycles), 68 °C for 3 min (1 cycle), hold at 4 °C.
2. Run the entire PCR product in a nucleic acid gel designed for DNA electrophoresis (*see* **Note 3**).
3. Using any gel extraction tool or a scalpel, excise the DNA band of the amplified gene of interest at the appropriate size, and place the gel plug in an Eppendorf tube.
4. Extract the PCR product using the gel extraction kit according to the manufacturer's instructions (*see* **Note 4**).
5. Digest the purified PCR product with endonucleases that recognize the restriction sites introduced at the 5′ and 3′ end of your gene (*see* **Note 2**).
6. Isolate the digested product by DNA electrophoresis (*see* **Note 3**).
7. Using any gel extraction tool or a scalpel, excise the DNA band of the digest, and place the gel plug in a microcentrifuge tube.
8. Extract the digested fragment using the gel extraction kit according to the manufacturer's instructions and eluting in 50 μL water (*see* **Note 4**).
9. To determine purity, run 10 μL of the gel-purified digest in a nucleic acid gel designed for DNA electrophoresis. A bright, clean band should be visible (*see* **Note 3**).

3.1.3 Ligation, Transformation, and Screening of Bacterial Clones

1. Perform a ligation reaction using the gel-purified digested vector, and insert with the DNA ligation kit according to manufacturer's instructions (*see* **Note 6**).
2. Using 1–3 μL of the ligation reaction, transform chemically competent *E. coli* in 250 μL SOC medium according to the manufacturer's instructions.

Fig. 1 (continued) To prepare the gene insert, two unique restriction sites are introduced through PCR amplification. Digestion at these sites must produce sticky ends that will be compatible and in proper orientation with those produced in the linearized vector. The PCR product is then gel-purified and digested with compatible restriction endonucleases. The digested gene insert is gel-purified and ligated to the gel-purified, linearized vector. Chemically competent *E. coli* are transformed with the ligation reaction mixture, and miniprep cultures of isolated colonies are grown to screen for the correct recombinant clone. The correct clone is scaled up, and plasmid DNA is extracted by performing a maxiprep purification. Sequencing verified identity of the resulting DNA

3. Plate 100 μL of the transformation mixture on one or more LB agar plates with 50 μg/mL kanamycin.
4. Invert plates and incubate overnight (approximately 18 h) at 37 °C.
5. Using a disposable inoculation loop, pick an isolated colony, and seed in 5 mL of LB broth with 50 μg/mL kanamycin in 15 mL conical tubes. Seed at least 5–10 cultures to screen for the correct clone.
6. Incubate cultures overnight (approximately 8–18 h) at 37 °C on a shaking incubator at 225 rpm.
7. Save 1 mL of each culture in labeled microcentrifuge tubes at 4 °C. The saved culture will be used for large-scale preparation when the correct recombinant clone is identified.
8. Isolate plasmid DNA from the remaining 4 mL of bacterial cultures using the miniprep kit according to the manufacturer's instructions.
9. Perform diagnostic restriction digests to screen plasmid minipreps (*see* **Note 7**).
   (a) Digest approximately 5 μL of each clone with the two restriction enzymes that generated the fragments used for the ligation step (*see* **Note 2**).
   (b) Run digested miniprep DNAs in a nucleic acid gel designed for DNA electrophoresis (*see* **Note 3**). Two bands should be visible under UV: the vector backbone (5.5 kbp) and the gene insert. The presence of two bands at their expected sizes indicates that the ligation and transformation reaction was successful.

3.1.4 Scale-Up of Recombinant Clone

1. Seed 1 mL of culture of the bacterial clone containing the correct recombinant plasmid (this culture was saved in **step 7** of Subheading 3.1.3) into 250 mL of LB broth with 50 μg/mL kanamycin in a 1 L Erlenmeyer flask.
2. Incubate the culture for 18–24 h at 37 °C in a shaking incubator at 225 rpm.
3. Isolate plasmid DNA from the bacterial culture using the plasmid maxi kit according to the manufacturer's instructions. The following modifications were made to enhance plasmid DNA purification:
4. After isopropanol precipitation and high-speed centrifugation of the eluted plasmid DNA, resuspend the pellet in 450 μL of ultrapure water in a microcentrifuge tube.
5. Add 50 μL (10%) of 3 M sodium acetate.

6. Add 1 mL (two volumes) of 100% ethanol to precipitate plasmid DNA. This precipitation step can also proceed at −20 or −80 °C overnight.
7. Precipitate DNA by centrifugation at maximum speed (20,800 × *g*) in a microcentrifuge at 4 °C for at least 10 min.
8. Carefully remove supernatant and wash the DNA pellet in 100% ethanol.
9. Spin the DNA at maximum speed (20,800 × *g*) in a microcentrifuge at 4 °C for 5 min.
10. Carefully remove supernatant, and let the DNA pellet air dry.
11. Resuspend the pellet in 500 μL of ultrapure water.
12. To determine quality and purity of the isolated plasmid, run 1 μL of the plasmid in a nucleic acid gel designed for DNA electrophoresis (*see* **Note 3**).
13. Sequence the DNA to verify that the recombinant clone contains the correct sequence.

3.2 Production of Ruc-Antigen

3.2.1 COS1 Cell Culture (See Note 8)

1. Follow the manufacturer's instructions if thawing COS1 cells from a frozen vial.
2. Cell cultures should be maintained as monolayers in T150 flasks using its complete growth medium.
3. Medium should be renewed every 3 or 4 days, and cell cultures should be split upon reaching confluency.

3.2.2 Transfection of Cell Culture with Recombinant Plasmid (See Note 8)

1. Plate COS1 cells at 1×10^6 cells per well of a 6-well cell culture plate in complete growth medium.
2. Incubate until cell monolayers are approximately 95% confluent. Incubation time may be 2 or 3 days until cells reach appropriate confluency.
3. When cells are ready for transfection, renew the culture with 2 mL fresh medium per well.
4. Lipofectamine LTX with Plus reagent is used for this transfection protocol according to the manufacturer's instructions with the following modifications outlined in **steps 5–8**.
5. DNA is diluted in 500 μL of Opti-Mem medium per well of a 6-well plate.
6. After addition of Plus reagent to diluted DNA, the solution is incubated for 15 min.
7. Lipofectamine LTX is directly added to the diluted DNA/Plus reagent solution, and the reaction is incubated for additional 30 min.
8. Five-hundred microliter of the transfection mixture is added to one well of a 6-well plate in a dropwise manner.

(a) Optimum transfections conditions may have to be determined empirically for each plasmid construct. The following conditions per well of a 6-well plate were used for recombinant plasmids encoding Ruc fused to norovirus antigens or the empty vector: 4 μg pRen2-VP1 DNA, 4 μL of Plus reagent, and 14 μL of Lipofectamine LTX in 500 μL of Opti-Mem medium. Five microgram pRen2-shell DNA, 5 μL of Plus reagent, and 11.25 μL of Lipofectamine LTX in 500 μL of Opti-Mem medium. 2.5 μg pRen2-protruding DNA, 2.5 μL of Plus reagent, and 11.25 μL of Lipofectamine LTX in 500 μL of Opti-Mem medium. Five microgram pRen2 DNA, 5 μL of Plus reagent, and 11.25 μL of Lipofectamine LTX in 500 μL of Opti-Mem medium.

9. After transfection, incubate cell culture for 48 h.

3.2.3 Harvesting of Ruc-Antigen

1. 48 h post transfection, wash cells once with 1× PBS.
2. Remove PBS, and add 1.4 mL of cold lysis buffer A to each 6-well plate, or approximately 235 μL per well, of transfected cells.
3. Scrape cells using a cell scraper from the surface of the plate, and split into two microcentrifuge tubes (approximately 700 μL into each tube) on ice.
4. Sonicate cells in tubes in an ultrasonic ice water bath, such as with a cup horn, at an amplification of 60% with a pulse of 5 s on and 5 s rest. Repeat for three cycles, making sure to keep the bath ice cold.
5. Clarify the crude lysates by centrifugation at 12,000 × *g* for 10 min at 4 °C.
6. Collect supernatant of cell lysates, containing the Ruc-antigen, and store in a clean microcentrifuge tube at −80 °C for up to a year.

3.3 LIPS Assay

3.3.1 Master Plate Preparation of Serum Samples

1. If testing a large number of samples, a master plate of sera can be prepared in advance (*see* **Note 9**).
2. Add 10 μL of serum sample to 90 μL of buffer A (1:10 dilution) to one well of a deep-well plate for a total of 100 μL of diluted sample.
3. Record the sample number and its corresponding location in the deep-well plate.
4. Seal the plate with a deep-well plate sealer or adhesive plate seals.
5. Store the master plate at 4 °C for up to a month as a ready-to-test solution for the LIPS assay (*see* **Note 10**).

3.3.2 Measuring Input Relative Luminometer Units (RLU) of Clarified Ruc-Antigen Lysates

1. To one well of a 96-well white, polystyrene, and opaque bottom microtiter plate, dilute 1 μL of Ruc-antigen lysate in 8 μL of 1× PBS for a total volume of 9 μL.
2. Measure RLUs emitted by Ruc-antigen lysate in a luminometer with a 50 μL injection of 1× coelenterazine substrate in *Renilla* assay buffer (*see* **Note 11**).
3. The output is RLUs per μL of Ruc-antigen lysate. Calculate microliters needed for 1×10^7 RLUs, which is the input amount of lysate used for each LIPS assay reaction (*see* **Note 12**).

3.3.3 LIPS Assay Procedure

A general schematic of the LIPS assay procedure is outlined in Fig. 2.

1. In a 96-well vinyl microtiter U-bottom plate, incubate Ruc-antigen and serum antibodies for 1 h at room temperature (*see* **Notes 10**, **12**, and **13**). This is achieved by adding 40 μL of buffer A, 10 μL of the 1:10 diluted sera from the master plate, and 10^7 RLUs of Ruc-antigen diluted into 50 μL buffer for a total of 100 μL of the antigen-antibody mixture.
2. During incubation of serum and Ruc-antigen lysate, dilute protein A/G plus UltraLink resin beads to a 30% suspension in 1× PBS in a microcentrifuge tube (*see* **Note 14**).

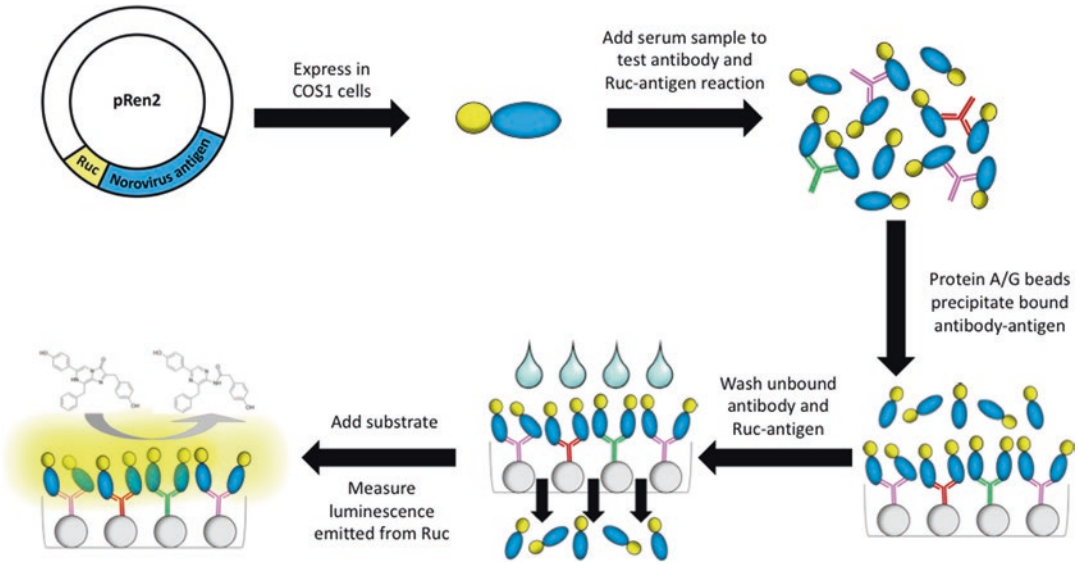


Fig. 2 LIPS assay schematic. Ruc fused to the antigen of interest is expressed in mammalian COS1 cells transfected with the corresponding recombinant plasmid. The fusion protein is harvested from crude cell lysates and added to sera for testing. Protein A/G beads in 96-well filter plates pull down antibody-antigen complexes from the serum-Ruc-antigen mixture. Unbound Ruc-antigens and other serum proteins are then washed and filtered through the plates. Coelenterazine substrate is added to the well, and only Ruc-antigen bound to precipitated antibodies will luminesce. Measurement of the luminescence will directly correlate with the level of antigen-specific antibodies in the serum sample

3. For each serum sample to be tested, add 5 μL of the 30% protein bead suspension to one well of a 96-well filter HTS plate (*see* **Note 15**).
4. Transfer the serum and Ruc-antigen lysate mixture to the protein A/G bead-coated filter plate.
5. Incubate for 1 h at room temperature.
6. Wash the filter plate on a vacuum manifold using 200 μL/well of buffer A once, 100 μL/well of buffer A seven times, and 100 μL/well of 1× PBS three times for a total of 11 washes (*see* **Notes 16** and **17**).
7. After the final wash, measure luminescence with a microplate luminometer using a 50 μL injection of 1× coelenterazine substrate in *Renilla* assay buffer from the *Renilla* luciferase assay system kit (*see* **Note 11**).

3.4 Analysis

Relative light units (RLUs) of serum samples tested in duplicate are averaged, and then the average RLU of the serum sample is corrected by calculating fold rise over the average background RLU. Background should be determined as RLUs measured from the LIPS assay of preimmunization or uninfected sera and the respective Ruc-antigen. Two or more independent experiments should be performed, and the resulting fold rise data are averaged to give mean fold rise in RLUs (Figs. 3 and 4) [15].

In Fig. 3, serum antibodies recognizing the major capsid protein VP1 or its individual structural domains, S or P, of a GII.4/2006b human norovirus strain were measured in two minipigs immunized with a cocktail of virus-like particles of several norovirus strains. Each respective time point included data of average RLUs of each sample. Minipigs were given booster doses at postimmunization weeks (PIW) 2, 4, and 9. Sera were collected before immunization (PIW 0) and at five time points (PIWs 2, 4, 7, 11, and 15) postimmunization. The mean fold rise in RLUs of antibodies increased over time against Ruc-VP1, Ruc-S, and Ruc-P, but not against Ruc alone, showing virus-specific antibody recognition. Antibody responses increased throughout the time course due to the multiple booster doses. These results show the LIPS assay can measure antibodies recognizing individual structural domains in addition to changes in the serum antibody response over time to a multivalent immunization.

The LIPS assay can also be used to study antibody responses to different viral strains (Fig. 4). In Fig. 4, Ruc-VP1 matching each component of the formulation used to immunize minipigs was used to assess antibody specificity in minipig sera at PIW 15. Each recombinant VP1 antigen was also used to analyze the reactivity of guinea pig hyperimmune serum raised against a GII.4/2004 human norovirus variant. A serologic response was detected against each Ruc-VP1 antigen in both the guinea pig monovalent and

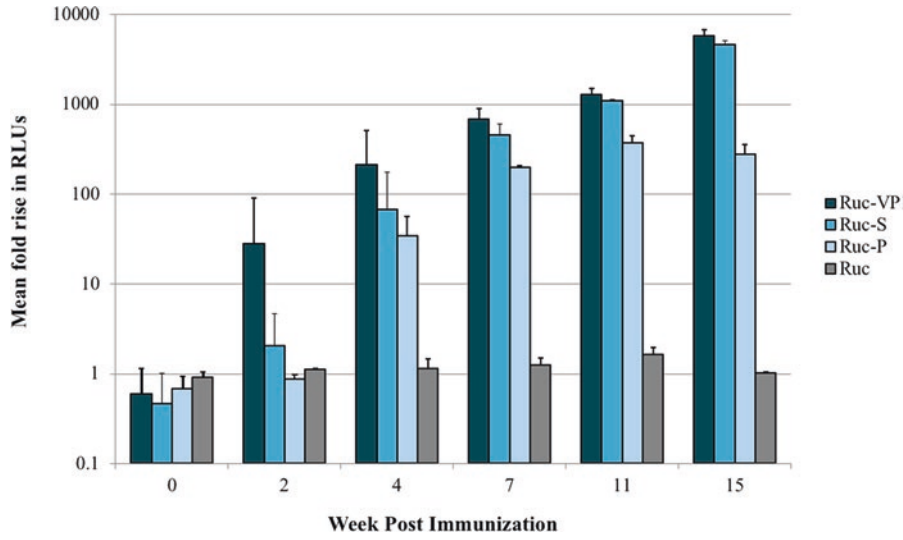


Fig. 3 LIPS analysis of minipig postimmunization sera. Human norovirus-specific antibody detection was presented as the mean fold rise in RLUs of two or more independent assay runs for two minipigs and plotted on a logarithmic scale. Background was determined as RLUs measured by the LIPS assay for negative control sera and the respective Ruc-antigen or Ruc alone as negative control antigen. Fold rise was calculated by dividing the average RLU of serum samples tested in duplicate by the average background. Error bars represent the standard error of the mean fold rise for two or more independent runs

minipig multivalent immunization sera, suggesting that the LIPS assay can measure both broadly reactive and strain-specific antibodies. In this data, a cutoff value to determine whether a sample was positive for antigen-specific antibodies was determined as the mean fold rise in RLUs of negative control sera plus three standard deviations.

4 Notes

1. Any buffers made in-house or general supplies and instruments purchased from different companies may be a suitable alternative to the materials listed in Subheading 2.
2. For restriction enzyme digestions, 10 U of enzyme are typically used to digest 1 μg of DNA in 1 h. Double digestions can be performed using high-fidelity restriction endonucleases and CutSmart® buffer such as that provided by New England BioLabs Inc. Otherwise, use the appropriate buffer recommended by the restriction enzyme source.
3. For nucleic acid gel electrophoresis, the gel should be run until the ladder is fully resolved (approximately 1 h at 100 V using a Bio-Rad mini-sub cell GT electrophoresis gel box). All samples are loaded alongside 8 μL of a 0.1 μg/μL DNA ladder on a 1% TAE agarose gel using 1× TAE running buffer for

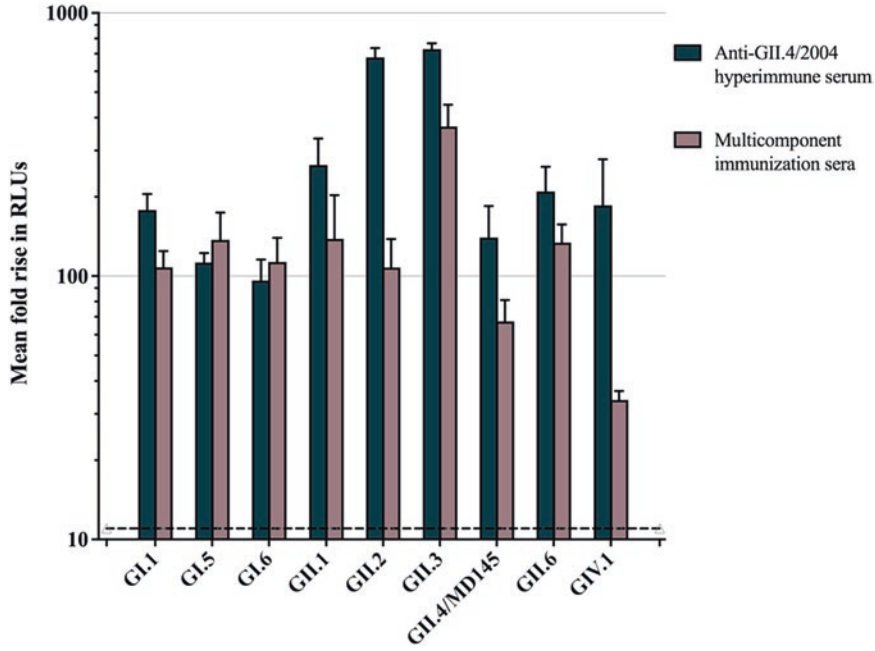


Fig. 4 LIPS analysis of postimmunization sera using Ruc-antigens of different human norovirus genogroups and genotypes. Human norovirus-specific serum antibodies against Ruc-VP1 of different strains were measured for hyperimmune guinea pig and multicomponent minipig immunization sera. Antigen-specific antibody detection was presented as the mean fold rise in RLUs of two or more independent runs and plotted on a logarithmic scale. Background was determined as RLUs measured by the LIPS assay for negative control preimmunization sera and each Ruc-VP1 antigen. Fold rise was calculated by dividing the average RLU of serum samples tested in duplicate by the average background. The cutoff value was calculated as the mean fold rise of mock-immunized minipig serum samples from multiple independent runs for each Ruc-antigen. The black, horizontal dashed line represents the cutoff value for positive human norovirus-specific antibody detection. Error bars represent the standard error of the mean fold rise for two or more independent runs

nucleic acid electrophoresis and 0.5 μg/mL ethidium bromide for staining.

4. To obtain more concentrated gel-purified DNA, the spin column can be incubated with less volume, such as 25 μL rather than 50 μL of ultrapure water, for 5 min prior to elution.
5. When designing oligonucleotide primers to introduce restriction sites into your gene insert, these restriction sites should not already be present in your gene. Generally, at least 6 bp should precede the recognition site of the restriction enzyme to allow proper cleavage, and endonucleases should generate sticky ends that will be compatible to those in the digested vector. It is important to design primers that will fuse the antigen-coding sequence in-frame and will also maintain proper orientation of the gene for unidirectional cloning.

6. For ligation reactions, a molar ratio of 1:3 vector to insert is generally used, taking into consideration the size of the vector and the insert DNA.
7. After isolating plasmid DNA from minipreps, clones can also be sequenced to screen for the correct recombinant prior to scaling up DNA for a maxiprep.
8. When handling mammalian cell culture, proper aseptic technique must be practiced at all times. Work is to be performed under a class II laminar flow hood specifically designed to prevent contamination in cell culture experiments.
9. Serum samples are tested in duplicate in the LIPS assay. Thus, when preparing the master plate of sera, it is helpful to prepare two wells side-by-side per serum sample for quick and efficient transfer to LIPS assay plates.
10. Prior to using a master plate of sera for testing with the LIPS assay, make sure to extensively shake the plate on an orbital shaker or mix samples well by pipetting up and down.
11. Optimum settings for the microplate luminometer may have to be determined empirically for each laboratory, depending on its brand or software used for calculations. For the settings used in this chapter, the plate was shaken for 2 s and had an integration read of 5 s for automatic injection of substrate. For manual injection, settings included shaking the plate for 2 s and an integration read of 0.2 s.
12. Relative light units (RLU) per μL of Ruc-antigen lysates must be calculated immediately prior to use in the LIPS assay as RLUs may slightly decrease after thawing stored lysates.
13. Ruc-antigen lysates that were stored at −80 °C should be thawed on ice prior to use.
14. Protein A/G beads must be vortexed thoroughly prior to dilution in a 30% PBS suspension.
15. If using the LIPS assay to test a large number of samples, the 30% protein bead suspension may be poured into a reagent reservoir, and a multichannel pipette may be used to transfer 5 μL to each well. However, the suspension must be mixed well prior to each transfer because the resin settles quickly.
16. An electronic multichannel pipettor is helpful for manual washing of the LIPS assay filter plate. On the other hand, a vacuum filtration system equipped with an automatic microplate washer can also be used instead of manual washing and filtration.
17. It is important to use a vacuum filter with hosing on the vacuum waste canister to prevent suction of fluids into the laboratory vacuum pump system.

Acknowledgments

This research was funded by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health. We thank Dr. Peter Burbelo for providing the pRen2 plasmid and Dr. Kim Green for helpful discussion of the manuscript.

References

1. Lanata CF, Fischer-Walker CL, Olascoaga AC, Torres CX, Aryee MJ, Black RE, Child Health Epidemiology Reference Group of the World Health Organization, UNICEF (2013) Global causes of diarrheal disease mortality in children <5 years of age: a systematic review. PLoS One 8(9):e72788. https://doi.org/10.1371/journal.pone.0072788
2. Green KY, Ando T, Balayan MS, Berke T, Clarke IN, Estes MK, Matson DO, Nakata S, Neill JD, Studdert MJ, Thiel HJ (2000) Taxonomy of the caliciviruses. J Infect Dis 181(Suppl 2):S322–S330. https://doi.org/10.1086/315591
3. Glass PJ, White LJ, Ball JM, Leparc-Goffart I, Hardy ME, Estes MK (2000) Norwalk virus open reading frame 3 encodes a minor structural protein. J Virol 74(14):6581–6591
4. Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK (1999) X-ray crystallographic structure of the Norwalk virus capsid. Science 286(5438):287–290
5. Lindesmith LC, Beltramello M, Donaldson EF, Corti D, Swanstrom J, Debbink K, Lanzavecchia A, Baric RS (2012) Immunogenetic mechanisms driving norovirus GII.4 antigenic variation. PLoS Pathog 8(5):e1002705. https://doi.org/10.1371/journal.ppat.1002705
6. Lindesmith LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, Vinje J, Baric RS (2008) Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med 5(2):e31. https://doi.org/10.1371/journal.pmed.0050031
7. Madore HP, Treanor JJ, Pray KA, Dolin R (1986) Enzyme-linked immunosorbent assays for Snow Mountain and Norwalk agents of viral gastroenteritis. J Clin Microbiol 24(3):456–459
8. Kapikian AZ, Greenberg HB, Cline WL, Kalica AR, Wyatt RG, James HD Jr, Lloyd NL, Chanock RM, Ryder RW, Kim HW (1978) Prevalence of antibody to the Norwalk agent by a newly developed immune adherence hemagglutination assay. J Med Virol 2(4):281–294
9. Hutson AM, Atmar RL, Marcus DM, Estes MK (2003) Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. J Virol 77(1):405–415
10. Hayashi Y, Ando T, Utagawa E, Sekine S, Okada S, Yabuuchi K, Miki T, Ohashi M (1989) Western blot (immunoblot) assay of small, round-structured virus associated with an acute gastroenteritis outbreak in Tokyo. J Clin Microbiol 27(8):1728–1733
11. Lewis DC, Lightfoot NF, Pether JV (1988) Solid-phase immune electron microscopy with human immunoglobulin M for serotyping of Norwalk-like viruses. J Clin Microbiol 26(5):938–942
12. Burbelo PD, Ching KH, Klimavicz CM, Iadarola MJ (2009) Antibody profiling by Luciferase Immunoprecipitation Systems (LIPS). J Vis Exp 32. https://doi.org/10.3791/1549
13. Burbelo PD, Bren KE, Ching KH, Gogineni ES, Kottilil S, Cohen JI, Kovacs JA, Iadarola MJ (2011) LIPS arrays for simultaneous detection of antibodies against partial and whole proteomes of HCV, HIV and EBV. Mol BioSyst 7(5):1453–1462. https://doi.org/10.1039/c0mb00342e
14. Burbelo PD, Hoshino Y, Leahy H, Krogmann T, Hornung RL, Iadarola MJ, Cohen JI (2009) Serological diagnosis of human herpes simplex virus type 1 and 2 infections by luciferase immunoprecipitation system assay. Clin Vaccine Immunol 16(3):366–371. https://doi.org/10.1128/CVI.00350-08
15. Tin CM, Yuan L, Dexter RJ, Parra GI, Bui T, Green KY, Sosnovtsev SV (2017) A Luciferase Immunoprecipitation System (LIPS) assay for profiling human norovirus antibodies. J Virol Methods 248:116–129. https://doi.org/10.1016/j.jviromet.2017.06.017



Chapter 9

Intact Mass Spectrometry Analysis of Immuno-Isolated Human Therapeutic Antibodies from Serum

Christie E. Delaney, John F. Kelly, Wen Ding, and Arsalan S. Haqqani

Abstract

Antibody-based therapeutics have emerged as novel class of biopharmaceuticals over the last couple of decades with the advancements made in production and downstream processing technologies. The structural diversity of therapeutic antibodies has also evolved with the development of bispecific (and multispecific) antibodies and antibody-drug conjugates. With increased structural complexities and multi-modularity, there is a need to demonstrate that the entire structure is stable in vivo and arriving at its target site in an intact form. Proving that antibodies reach their target site unscathed is a challenging but essential step for showing effective delivery as well as showing whether failure in efficacy (if any) was related to its in vivo instability. This chapter describes a method for highly specific immuno-isolation followed by intact mass spectrometry of human Fc-containing antibody from serum of rats dosed with the antibody. The method provides an opportunity for evaluating antibody stability in the physiological environment by providing accurate validation of its molecular mass in vivo, as well as the potential to identify breakdown products.

Key words Serum stability, Breakdown products, Affinity purification, IgSelect, Intact LC-MS

1 Introduction

The therapeutic antibody market has continued to grow since the first antibody was marketed in 1987, and currently there are over 80 monoclonal antibody therapeutics approved for use in North America and Europe [1–3]. The structure of the therapeutic antibody has also evolved with the development of technologies for adding additional domains to the already multi-modular immunoglobulin structure, giving the antibody multiple functionalities [4]. Through exploiting this technology, more than 60 different multifunctional antibody formats of different degrees of complexities have been created [5]. These have typically been achieved through combination of various domains (e.g., VH, VL, Fab, scFv, V_HH) and/or by creating bispecific antibodies through strategic heterodimerization of two antibody fragment "halves" [5–8]. In addition, conjugation of cytotoxic drugs to antibodies to create

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_9, © Springer Science+Business Media, LLC, part of Springer Nature 2019

antibody-drug conjugates (ADCs) has also emerged as a novel class of anticancer treatment agents that combine the selectivity of targeted treatment with the potency of chemotherapeutics [9, 10]. However, the increase in structural complexity, involving additional domains and drug conjugations, has also led to challenges in producing a stable, fully intact antibody structure and validating its stability in vivo [4, 11–13].

Several methods can be used to validate whether a newly produced antibody or ADC has a fully intact structure [12]. UV/vis spectroscopy and hydrophobic interaction chromatography (HIC) are useful techniques for determining drug-antibody ratio (DAR) of an ADC. However, UV/vis will only give an indication of the average DAR, and HIC is often hindered by heterogeneous mixtures of conjugated species [4, 12]. The advantages of liquid chromatography coupled with intact protein mass spectrometry (LC-MS) are that analysis is quick and usually little sample is required, while useful mass and structural information can be obtained [14]. The field of LC-MS continues to grow, with greatly improved instrumentation and methods, and has found wide application in the study of monoclonal antibodies and antibody-based therapeutics. In addition to sample identification, intact protein MS can provide relative abundance information for the different protein species present in a heterogeneous mixture [15]. Antibodies, biotherapeutics, and ADCs can be analyzed by intact mass LC-MS for physical stability, DAR, sequence determination, presence of impurities, as well as the glycosylation profile [12, 14, 15].

Antibodies and ADCs can become unstable in vivo for a variety of reasons [4, 11–13, 16], and their stability has traditionally been assessed by two methods: enzyme-linked immunosorbent assay (ELISA) and liquid chromatography tandem-mass spectrometry (LC-MS/MS). Although ELISA can provide valuable information regarding the ability of the biotherapeutic product to bind target in serum, it might not provide evidence of whether degradation or "clipping" of the therapeutic arm from the antibody has occurred. The signal from antibodies can be non-specific, and small modifications to the therapeutic arm or ADC may not be detected. In addition, ELISA requires multiple assay formats in order to achieve a full understanding of the in vivo transformation of the biotherapeutic product [17]. LC-MS/MS and/or selective reaction monitoring (SRM) techniques can be useful for detecting the released drug payload or the therapeutic arm, but does not provide information about the overall structural integrity of the biotherapeutic. Also, given the frequent difficulties of predicting all of the important degradation products of a biotherapeutic or ADC, the targeted LC-MS/MS approach may not be ideal [15].

Affinity capture has been used in combination with proteomics, SRM, MS, and HIC to isolate biotherapeutics from serum and

plasma for post-in vivo biophysical characterization to monitor for degradation, such as cleavage of the therapeutic arm or drug payload [12, 15]. When coupled with immuno-isolation, intact protein LC-MS becomes a powerful tool for determining serum stability of a biotherapeutic product. To utilize intact mass LC-MS effectively, the biological samples need to be of high purity with low contamination from endogenous host proteins. Therefore, immuno-isolation when developed and applied properly can be an excellent strategy for purifying the protein from serum prior to analysis by mass spectrometry [15].

This chapter provides a protocol for immuno-isolation of a human Fc (hFc)-containing antibody from rodent serum followed by intact protein LC-MS analysis. Various methods, including protein A and anti-hFc conjugated to magnetic beads, were initially tested for the immuno-isolation method. However, the affinity capture technique with the highest yield and purest results was obtained with an alkali-stable resin with a ligand specific for the hFc fragment, called IgSelect (GE Healthcare). Modifying a method investigated for antibody purification by Ishihara et al. [18], an hFc-containing antibody was immuno-isolated using IgSelect from serum collected from an intravenously injected rat (Fig. 1). The antibody was then analyzed by intact protein LC-MS to determine its molecular mass and structure, as a measure of its stability post-serum exposure in vivo. The method has further potential for immuno-isolation of antibody-based therapeutics from other matrices, such as the cerebrospinal fluid and brain.

2 Materials

2.1 Antibody of Interest

1. The antibody of interest must contain an hFc to be immuno-isolated by IgSelect (*see* **Note 1**). As an example, the antibody used for LC-MS following immuno-isolation is a bivalent 83,780 Da FC5-hFc-Pep antibody, comprised of a single domain antibody, FC5 [19], fused to an hFc fragment and a therapeutic peptide. An antibody lacking the therapeutic peptide, bivalent 78,360 Da FC5-hFc antibody, was used as a control.

2.2 Rodent Serum

1. At least 55 μL of test or spiked serum is required for each immuno-isolation and LC-MS analysis.
2. Test serum: The test serum used here was collected from the tail vein of male Wistar rats 6 h after administration of intravenous injection of 125 nmol/kg of FC5-hFc-Pep antibody. The serum was stored at −80 °C in glass vials until use (*see* **Note 2**).

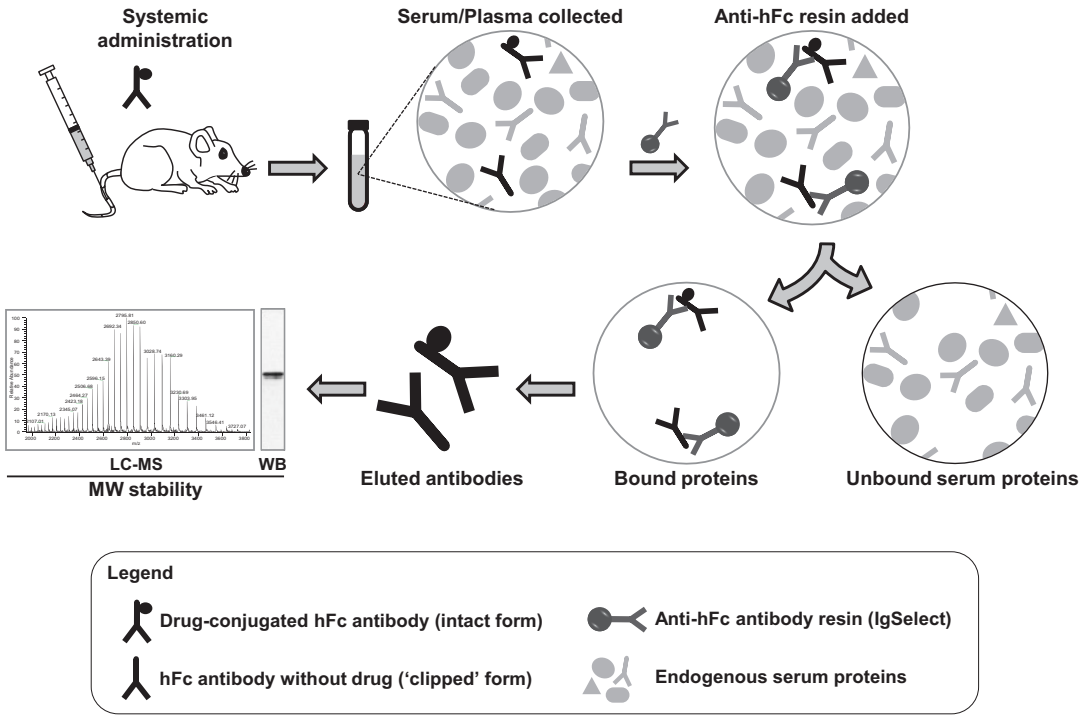


Fig. 1 A schematic showing the overview of the immuno-isolation procedure. An hFc-containing therapeutic antibody or ADC is administered to animals, and test serum samples are collected at various time points. The test serum may contain intact antibody, "unstable" antibody fragment ("clipped" form), and endogenous serum proteins. Anti-hFc IgSelect beads are added to the test serum to specifically bind the intact and "unstable" antibodies. Unbound serum proteins are removed, and IgSelect-bound proteins are eluted. The eluted proteins can be analyzed by LC-MS, western blotting (WB), and/or other methods to evaluate MW stability of the antibodies

3. Spiked serum: Spiked serum is used as a positive control for optimizing the immuno-isolation procedure. FC5-hFc-Pep antibody was spiked in control serum (*see* **Note 3**). Serum from naïve animals or vehicle-injected animals can be used as control serum. Alternatively, control serum from similar strain of animals can be purchased from a commercial source (e.g., BioIVT, USA).

2.3 IgSelect Immuno-Isolation

1. IgSelect affinity resin.
2. 1.5 mL low-retention micro centrifuge tubes.
3. Benchtop centrifuge: Eppendorf Centrifuge 5451D.
4. Vortex Genie II, fitted with a multi-tube holder.
5. Milli-Q water: Milli-Q® Ultrapure MS-grade water.
6. Binding buffer: 0.1 M Tris–HCl, 0.15 M NaCl, pH 7.2.
7. Elution buffer: 0.1 M glycine–HCl, pH 2.2.

8. Laemmli buffer (Bio-Rad Laboratories): 1×, diluted with ddH_2O.
9. 1 N sodium hydroxide.

2.4 LC-Intact Protein Mass Spectrometry

1. HPLC system such as the Agilent 1100 series equipped with a binary pump, autosampler, and online degasser and capable of performing reversed-phase liquid chromatography (RPLC) at 3 mL/min.
2. A high-resolution mass spectrometer such as the LTQ-Orbitrap XL™ mass spectrometer (Thermo Fisher) equipped with an Ion Max electrospray source (32-G needle). The instrument is calibrated with CsI and tuned for large protein (>50 kDa) analysis.
3. Poros R2 column, 2.1 × 30 mm (Applied Biosystems).
4. A stainless steel tee that is installed post-column to direct 100 µL/min of the column eluate to the electrospray ion source. The remaining flow is sent to waste or to a UV detector. Column oven or jacketed heater that can reach 80 °C.
5. Thermostated oven to heat the mobile phase prior to entering the column.
6. Solvent A: 0.1% (v/v) formic acid in HPLC grade water (e.g., Optima LC-MS Water).
7. Solvent B: 100% (v/v) acetonitrile (e.g., Omnisolv Acetonitrile).

2.5 Software and Databases for Data Analyses

1. Xcalibur version 2.2, including the Qual Browser module (Thermo Fisher).
2. MassLynx©, version 4.1, including the Databridge and MaxEnt 1 algorithms (Waters Ltd).

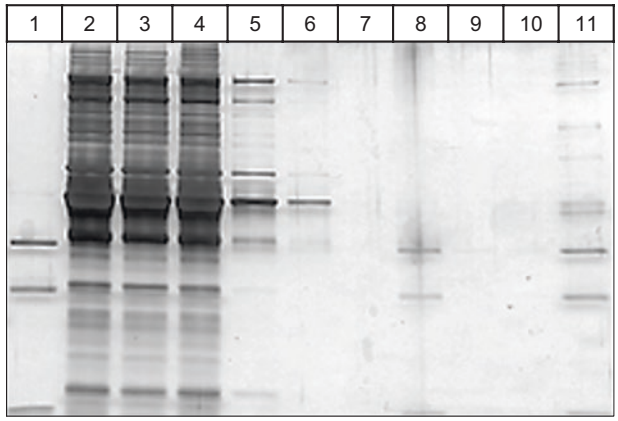
3 Methods

3.1 Immuno-Isolation with IgSelect Beads

The first step of the immuno-isolation procedure is to prepare the anti-hFc beads. A schematic showing the overview of the immuno-isolation procedure is shown in Fig. 1. Briefly, test sera containing the antibody are collected from dosed rodents and incubated with anti-hFc IgSelect beads (*see* **Note 4**). The unbound proteins are removed and the anti-hFc bound proteins are eluted. The eluate can be analyzed by various methods to determine protein stability, including intact LC-MS and gel-based methods (*see* **Note 5**). An example of SDS-PAGE showing immuno-isolation of an antibody from serum is shown in Fig. 2.

3.1.1 Preparation of the Beads

1. Transfer 100 µL of IgSelect slurry into a 1.5 mL tube for each serum sample to be analyzed (*see* **Note 6**).



1. hFc-containing antibody
2. Test serum
3. Unbound proteins
4. Wash 1
5. Wash 2
6. Wash 3
7. Wash 4
8. Elution 1
9. Elution 2
10. Elution 3
11. Un-eluted proteins

Fig. 2 SDS-PAGE of an hFc-containing antibody immuno-isolated from serum using IgSelect beads. A 100 kDa multi-domain antibody (lane 1) shows multiple bands on the reduced SDS-PAGE. A 6 h test serum from rats dosed with the same antibody was used as a starting material (lane 2). The test serum was incubated with IgSelect beads and unbound proteins (lane 3) were removed. The beads were washed four times (lanes 4–7), and the antibody was eluted in three fractions (lanes 8–10). Un-eluted proteins on the beads were stripped using Laemmli buffer (lane 11)

2. Centrifuge the beads at 60 × *g* (*see* **Note 7**).
3. Remove and discard the supernatant containing storage buffer.
4. Add 400 µL of binding buffer to the beads, and gently tap the tube to mix (*see* **Note 8**).
5. Invert the tube several times to wash the beads.
6. Centrifuge at 60 × *g* and discard the supernatant.
7. Repeat **steps 4–6** three times to thoroughly wash the beads for a total of four washes.

3.1.2 Binding Antibodies to Beads

1. To a fresh 1.5 mL Eppendorf tube, add 55.0 µL of the spiked or test serum.
2. Add 50 µL of binding buffer and vortex (*see* **Note 9**).
3. Remove and save a 5 µL aliquot as the "Input" (*see* **Note 10**).
4. Transfer the remaining 100 µL of the serum/binding buffer mixture to the 100 µL of prepared IgSelect beads from Subheading 3.1.1.
5. Tap gently to mix the serum/bead slurry thoroughly.
6. Incubate for 1 h at room temperature, at the lowest speed setting on a Vortex Genie, fitted with a tube holder (*see* **Note 11**).
7. Centrifuge at 60 × *g* for 1 min.
8. Remove the supernatant and save it as "Unbound serum proteins" (*see* **Note 10**).

3.1.3 Washing the Beads

1. To wash the beads, add 400 µL of binding buffer and tap gently to mix.
2. Vortex for 5 min at the lowest speed setting on the vortex (*see* **Note 12**).
3. Centrifuge at 60 × *g* for 1 min.
4. Save the supernatant as "Wash" (*see* **Note 10**).
5. Repeat **steps 1–4** three times to wash the beads for a total of four times, and save all of the washes in separate, labeled tubes as "Wash" 1, 2, 3, and 4 (*see* **Note 13**).

3.1.4 Eluting Antibodies from the Beads

1. Add 50 µL of elution buffer and tap gently to mix (*see* **Note 14**).
2. Incubate for 5 min at room temperature on the lowest setting of the vortex.
3. Centrifuge at 60 × *g* for 1 min.
4. Remove the supernatant and save the eluted fraction in a 1.5 mL tube labeled "Eluted proteins" (*see* **Note 10**).
5. Add 5.0 µL of 1 N NaOH to the eluted fraction to neutralize the pH (*see* **Note 15**).
6. Repeat the elution at least twice more (**steps 1–5**), and collect the eluted fractions separately as "Eluted proteins" 1, 2, and 3 (*see* **Note 16**).

3.1.5 Stripping Un-eluted Proteins from the Beads

1. To strip un-eluted proteins, add 50 µL of 1× Laemmli buffer to the beads (*see* **Note 17**).
2. Centrifuge at 60 × *g* for 1 min.
3. Transfer the supernatant to a fresh tube labeled "Un-eluted proteins" (*see* **Note 10**).
4. Add 4.0 µL of 1 N NaOH to the "Un-eluted proteins" to neutralize the pH (*see* **Note 18**).
5. The samples may be stored at −20 °C for up to 4 weeks or analyzed directly by various methods, including intact LC-MS (Subheading 3.2) and gel-based methods (Fig. 2).

3.2 LC-Intact Mass Spectrometry Analysis

3.2.1 LC-MS Setup

1. The Poros R2 column is equilibrated at 20% solvent B at a flow rate of 3 mL/min prior to injecting the sample. The column oven temperature is set to 82.5 °C, and the mobile phase is pre-warmed in a separate temperature-controlled oven set to 80 °C (*see* **Note 19**).
2. The FTMS resolution of the LTQ-Orbitrap XL mass spectrometer is set to 7500. The electrospray voltage is set to 3.8 kV, the capillary voltage to 40 V, and the tube lens to 170 V (*see* **Note 20**).

3.2.2 Sample Analysis

1. Centrifuge the eluted protein at 20,000 × *g* for 5 min, and transfer the supernatant to a new 1.5 mL tube (*see* **Note 21**).
2. Load up to 5 μg of protein sample onto the Poros column (*see* **Note 22**).
3. The RPLC gradient is as follows: solvent B is (1) maintained at 20% for the first 3 min, (2) increased from 20% to 90% from 3 to 6 min, (3) decreased from 90% to 20% for 6–7 min, and (4) maintained at 20% from 7 to 9 min. Hold at 20% for 2 min (*see* **Note 23**).
4. The eluate from the column passes through the divert valve on the LTQ-Orbitrap XL before going to the splitting tee. To avoid contamination of the ESI source, the valve is activated at the beginning of each run for 1.5 min in order to divert the salt in the samples to waste.
5. MS spectra are acquired from m/z 400 to 4000. A start delay of 1.5 min ensures that there is no spectral acquisition when the divert valve is activated.
6. The LC-MS file is opened in Qual Browser (Xcaliber™), and the MS spectra acquired across the protein peak are summed. The resulting spectrum is exported to Excel (export: clipboard (exact mass)). The file header (top 8 lines) is deleted, and the spreadsheet is saved as a tab delimited file (.txt).
7. Use Databridge (MassLynx©) to convert the text file to a MassLynx compatible format. Select the text file to be converted, and identify the directory in which the converted file will be saved. Open the Options tab, and set the source file as "ASCII" and the target file as "MassLynx." To perform the conversion, click on the Convert tab, select the "Continuum" option, select "Electrospray+" as the ion mode, and press "OK."
8. Open the converted mass spectrum in MassLynx using the Spectrum module. Set the m/z range to cover the multiply charged ion envelope of the protein. Open MaxEnt1, which is located under the Process icon on the menu bar. Set the output mass range (i.e., 50,000–160,000 Da for full-sized antibodies) and resolution (i.e., 1.00 Da), the damage model (typically Uniform Gaussian, 1.00 Da width at half-height), the minimum intensity ratios (60% for both the left and right works well for full-sized antibodies), and the number of iterations (typically 10–15). Press "OK," and MaxEnt1 will attempt to generate a molecular weight profile from the mass spectral data (*see* **Note 24**).

3.2.3 An LC-MS Example

An example of LC-MS analysis of an immuno-isolated antibody from test serum is shown in Fig. 3. Briefly, a control serum was spiked with FC5-hFc-Pep antibody and first used to optimize the

immuno-isolation procedure (Subheading 3.1). A 6 h test serum sample from rat dosed with 125 nmol/kg of FC5-hFc-Pep antibody was then immuno-isolated using the procedure and analyzed by intact LC-MS (Fig. 3). As controls, FC5-hFc-Pep antibody and an FC5-hFc antibody (lacking the therapeutic peptide) were also analyzed by intact LC-MS. As seen in Fig. 3, the predominant mass detected from LC-MS analysis of FC5-hFc-Pep antibody purified from the test serum sample is 83,754 Da, which corresponds to the theoretical mass of the fully intact antibody (83,760 Da). This mass is clearly distinguishable from the 78,446 Da mass observed for the FC5-hFc antibody without the therapeutic peptide (Fig. 3c). Hence, this method has a potential to identify some of the unstable products or small modifications of the antibody/ADC if the hFc epitope is still intact.

3.3 Conclusions

Antibody-based therapeutics and ADCs have emerged as novel classes of biopharmaceuticals over the last couple of decades, and as their structures become more complex, additional needs arose to demonstrate that the final product is stable and fully intact in the physiological environment. LC-MS is currently the most accurate method for estimating the molecular mass of the intact antibody. Combining this method with a highly selective purification strategy has made it possible to tackle the challenge of more accurately validating the antibody molecular mass in vivo and hence demonstrating its stability. The method described here can be used to estimate the molecular mass of the antibody in vivo and compare it with the molecular mass of the starting product. The method also has a potential to be semiquantitative as well as the ability to identify some of the unstable breakdown products of the antibody or ADCs (*see* **Notes 25** and **26**). The method has further potential for analysis in other matrices, such as the cerebrospinal fluid and brain.

4 Notes

1. The method described here is for evaluation of an antibody-containing hFc domain. If an ADC is being tested, evaluation of both unconjugated antibody and ADC is recommended. If the antibody contains mouse Fc or other domains, alternative secondary antibodies should be used and optimized.
2. Typically serum to be tested is collected from animals dosed with the antibody of interest. The concentration of the antibody in test sample is critical. Concentrations below 0.05 μM should be avoided as the immuno-isolation may not be reproducible. Since the concentration is dependent on multiple factors (serum half-life, collection time, administered dose), it is

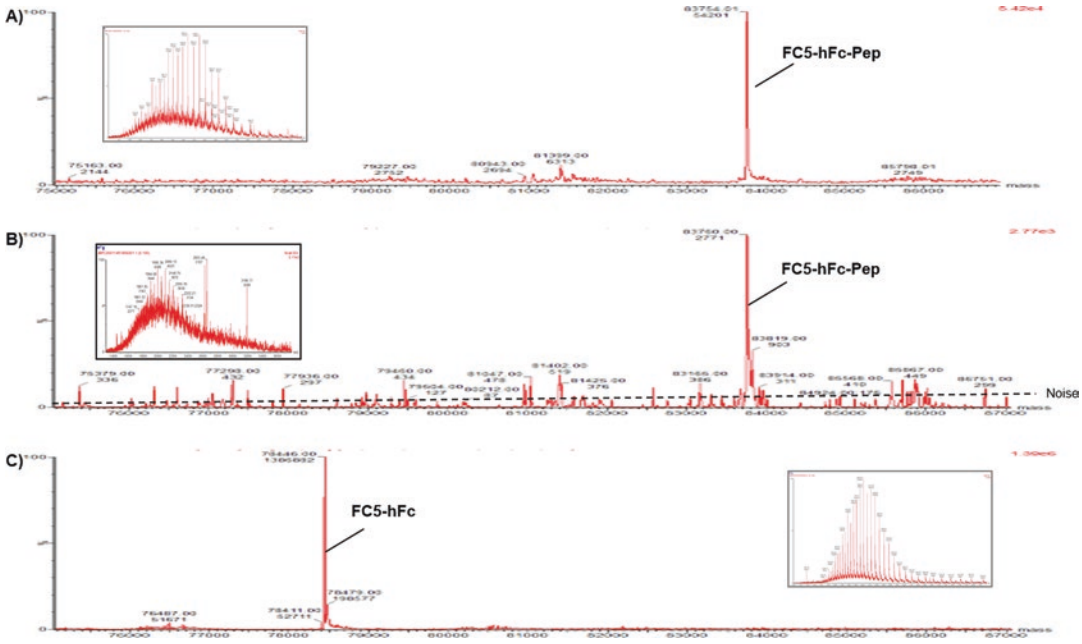


Fig. 3 LC-MS analysis of a serum-purified antibody. (**a**) A de-convoluted LC-MS spectrum of FC5-hFc-Pep antibody (theoretical MW 83,780 Da). Inset shows the original convoluted spectrum. The observed MW is 83,754 Da. (**b**) A de-convoluted LC-MS spectrum of the FC5-hFc-Pep antibody immuno-isolated from a 6 h test serum using IgSelect. Inset shows the original convoluted spectrum. The observed MW is 83,760 Da suggesting that the predominant form in serum is still intact. The test serum was from rats dosed with 125 nmol/kg of the FC5-hFc-Pep antibody. (**c**) A de-convoluted LC-MS spectrum of FC5-hFc antibody (theoretical MW 78,360 Da). Inset shows the original convoluted spectrum. The observed MW is 78,446 Da

recommended that it should be estimated using alternative methods prior to this experiment (such as SRM or ELISA).

3. As a positive control, control serum is mixed with known concentration of the antibody of interest, typically within physiological range. For example, FC5-hFc-Pep antibody administered at a 125 nmol/kg dose in rats typically gives a concentration range of 1.0–1.5 μM at 6 h. As mentioned in **Note 2**, low concentrations should be avoided.
4. It is highly recommended that the immuno-isolation procedure should first be tested using spiked serum to allow optimization of the protocol.
5. It is recommended that, in addition to LC-MS, other methods should be tested for evaluating protein MW, such as gel-based methods (SDS-PAGE and western blotting). Although they do not provide exact MWs and are semiquantitative, gel-based methods could be helpful in quickly evaluating the progress of the immuno-isolation procedure and testing antibody structural instability as determined by MW changes (e.g., deconju-

gation of ADCs), especially if also run under nonreducing conditions.

6. The agarose beads are delicate and must be handled with care. To pipette a homogenous volume of the IgSelect resin and avoid damage to the IgSelect agarose beads, cut the tip off of the pipette tip to increase the size of the opening. Avoid rigorous vortexing and mixing. Swirl the IgSelect to mix, pipette up and down gently to thoroughly resuspend the resin, and immediately pipette the desired volume, as it will settle quickly.
7. The use of a higher centrifugal speed than 60 × *g* might lead to rupture of the beads.
8. Due to the incompatibility of detergents with mass spectrometry, no detergents are used in the buffers for the immuno-isolation method, except for the purpose of stripping the "Un-eluted proteins" bound to the beads after the elution (Subheading 3.1.5).
9. It was determined that an approximate 1:1 ratio of serum to binding buffer was necessary for optimal binding in subsequent steps. Too little binding buffer led to poor binding of the antibody to the beads; however, sufficient serum must be used in order to detect the antibody.
10. Throughout the immuno-isolation procedure, it is highly recommended to collect and save the "Input," "Unbound serum proteins," "Washes," and "Un-eluted proteins," in addition to the "Elutions" for quality control purposes. To determine the purity and efficiency of the immuno-isolation, add Laemmli buffer, load an equivalent amount of each sample on SDS-PAGE, and visualize the proteins by silver staining (*see* Fig. 2) or by western blotting with an anti-hFc antibody.
11. Although an end-over-end rotator is traditionally used for immuno-isolation, it was found to be suboptimal for the current method due to the small volume of sample. For in vivo rat and mouse samples, serum is often limited. If sample is not limited, the volume of both sample and IgSelect can be increased. A traditional vortex at the lowest speed setting and fitted with a tube holder did not damage the beads significantly. Another option would be to use an incubator with a shake function at room temperature, such as the Benchmark Multi-Therm H5000 set to 200 rpm. It should be noted that using a vortex, even at the lowest setting, may lead to some losses of the delicate agarose beads, but the aim of this method is inherently qualitative and not quantitative.
12. To wash the beads and remove non-specifically bound antibody, it is necessary to shake the samples on the lowest speed

setting on the vortex. As discussed above (*see* **Note 11**), even using a vortex at lowest setting may damage the agarose beads.

13. For quality control purposes, it is recommended to save the washes separately. Refer to **Note 10**. Run the washes side-by-side on SDS-PAGE and subsequently western blot with anti-Fc or silver stain to determine the efficacy of the washing steps and the purity of the immuno-isolation.
14. Typically glycine is used to elute antibodies from the beads under acidic conditions. However, if a large amount of antibody is still not being eluted (Fig. 2), alternative acidic or basic elution buffers (e.g., citrate, diethylamine, and borate) may be tested to find the most optimal elution buffer for the antibody of interest.
15. Neutralization of the pH is only necessary for the purposes of SDS-PAGE. If desired, an aliquot for SDS-PAGE could be removed and neutralized.
16. The first eluted fraction is likely to be of highest purity; therefore, it is not recommended to pool the eluted fractions.
17. The bromophenol blue in the Laemmli buffer will turn its color from blue to yellow upon addition to the beads. Laemmli buffer without ß-mercaptoethanol is used to strip the un-eluted proteins from the beads. This is to ensure proteins (especially antibodies) stay in their native un-reduced form, which allows evaluation of the full structure of the antibody by SDS-PAGE (e.g., dimerization of the two halves of the bispecific antibodies).
18. Upon addition of NaOH, the bromophenol will turn back to blue color.
19. It was found that lower temperatures tend to give poorer peak shape, especially for larger proteins such as monoclonal antibodies.
20. These voltages are typical settings used for the LTQ-Orbitrap XL in our lab. These values may need to be adjusted depending on the instrument used.
21. It is important that the eluted protein sample is free of any beads or debris before injecting into the LC system in order to avoid any blockages.
22. Although the recommended injection amount is 5 μg of protein, it is possible to produce good signal with significantly less protein. If better sensitivity is required, then it is recommended to use a column designed for lower flow and that does not require a post-column flow splitter.
23. This RPLC gradient is fast and designed for effective desalting and high sample throughput. The Poros R2 column is an excellent fit for this purpose, but it is not intended to provide

good separation of protein mixtures. If resolution is a higher priority than sample throughput, then it is recommended to use a different column system such as the ProSwift C4 RO-5H columns from Thermo Fisher Scientific.

24. It was found that MaxEnt1 maintains the ratio of the ions in the spectrum better than other programs previously used for generating molecular weight profiles. The MaxEnt1 setting will need to be modified depending on the molecular weight of the proteins under investigation as well as the quality and complexity of the original mass spectrum.
25. The immuno-isolation LC-MS method has a potential to be a quantitative method. Briefly, a calibration curve can be prepared by spiking a range of physiological concentrations of the antibody of interest in serum to generate calibration standards. Each standard is then immuno-isolated, analyzed by LC-MS, and the integrated intensities are calculated. The intensities are then used for generation of a calibration curve from which the concentration of the test sample can be estimated.
26. The immuno-isolation LC-MS method also has a potential to identify some of the unstable products or small modifications of the antibody/ADC if the hFc epitope is still intact. An example is when the conjugate of an ADC is "clipped." In this case, running the unconjugated antibody and the ADC as controls would allow identification of the "clipped" forms in test serum since the latter will typically have molecular mass between the two controls.

References

1. Ecker DM, Jones SD, Levine HL (2015) The therapeutic monoclonal antibody market. MAbs 7:9–14
2. Schiel JE, Mire-Sluis A, Davis D (2014) Monoclonal antibody therapeutics: the need for biopharmaceutical reference materials. In: Schiel JE, Davis DL, Borisov OV (eds) State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization volume 1. Monoclonal antibody therapeutics: structure, function, and regulatory space. Am Chem Soc, Washington, DC, pp 1–34
3. Drugs@FDA: FDA approved drug products. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm. Accessed 10 Jul 2018
4. McCombs JR, Owen SC (2015) Antibody drug conjugates: design and selection of linker, payload and conjugation chemistry. AAPS J 17:339–351
5. Spiess C, Merchant M, Huang A, Zheng Z, Yang N-Y, Peng J, Ellerman D, Shatz W, Reilly D, Yansura DG, Scheer JM (2013) Bispecific antibodies with natural architecture produced by co-culture of bacteria expressing two distinct half-antibodies. Nat Biotechnol 31:753–758
6. Stanimirovic D, Kemmerich K, Haqqani AS, Farrington GK (2014) Engineering and pharmacology of blood-brain barrier-permeable bispecific antibodies. Adv Pharmacol 71: 301–335
7. Wickramasinghe D (2013) Tumor and T cell engagement by BiTE. Discov Med 16: 149–152
8. Chames P, Baty D (2009) Bispecific antibodies for cancer therapy: the light at the end of the tunnel? MAbs 1:539–547
9. Beck A, Wurch T, Bailly C, Corvaia N (2010) Strategies and challenges for the next generation of therapeutic antibodies. Nat Rev Immunol 10:345–352

10. Beck A, Goetsch L, Dumontet C, Corvaïa N (2017) Strategies and challenges for the next generation of antibody-drug conjugates. Nat Rev Drug Discov 16:315–337
11. Razinkov VI, Treuheit MJ, Becker GW (2015) Accelerated formulation development of monoclonal antibodies (mAbs) and mAb-based modalities. J Biomol Screen 20:468–483
12. Ross PL, Wolfe JL (2016) Physical and chemical stability of antibody drug conjugates: current status. J Pharm Sci 105:391–397
13. Correia IR (2010) Stability of IgG isotypes in serum. MAbs 2:221–232
14. Zhang H, Cui W, Gross ML (2014) Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. FEBS Lett 588:308–317
15. Xu K, Liu L, Saad OM, Baudys J, Williams L, Leipold D, Shen B, Raab H, Junutula JR, Kim A, Kaur S (2011) Characterization of intact antibody–drug conjugates from plasma/serum in vivo by affinity capture capillary liquid chromatography–mass spectrometry. Anal Biochem 412:56–66
16. Diamantis N, Banerji U (2016) Antibody-drug conjugates—an emerging class of cancer treatment. Br J Cancer 114:362–367
17. Sauerborn M, van Dongen W (2014) Practical considerations for the pharmacokinetic and immunogenic assessment of antibody–drug conjugates. BioDrugs 28:383–391
18. Ishihara T, Nakajima N, Kadoya T (2010) Evaluation of new affinity chromatography resins for polyclonal, oligoclonal and monoclonal antibody pharmaceuticals. J Chromatogr B 878:2141–2144
19. Tanha J, Muruganandam A, Stanimirovic D (2003) Phage display technology for identifying specific antigens on brain endothelial cells. Methods Mol Med 89:435–449



Chapter 10

Antigen Identification for Cell-Binding Antibodies Using Ligand-Directed Crosslinking and Biotin Transfer

Tammy-Lynn Tremblay and Jennifer J. Hill

Abstract

Panning approaches using antibody libraries often result in the isolation of antibodies that bind to cells through an unknown cellular receptor. Here, we describe a protocol that uses ligand-directed crosslinking with the aminooxy-sulfhydryl-biotin (ASB) trifunctional crosslinker followed by a proteomic analysis to identify the cellular receptors for orphan ligands. We describe the synthesis of the ASB crosslinker, labelling of the ligand with ASB, and cell binding of the labelled ligands. Next, biotin affinity purification and trypsin digestion of cell surface proteins that have been crosslinked by ASB are described. Lastly, several hints and tips to improve the proteomic analysis for these types of samples are provided.

Key words Receptor identification, Crosslinking, Proteomics, Mass spectrometry, Antibodies

1 Introduction

Over the past decade, technological advances have improved the ability to isolate monoclonal antibodies from humans and from synthetic human antibody libraries [1]. Many of these new human libraries rely on cell-based panning or isolation approaches in order to identify antibodies that recognize antigens with the desired cell-binding and functional characteristics [2–5]. Thus, these approaches often result in the isolation of an antibody that recognizes an unidentified antigen on a cell surface. The subsequent identification of the antigen can be complicated by the highly hydrophobic nature of cell membrane proteins as well as difficulty expressing or isolating these antigens outside of the cell membrane.

In this chapter, we describe a detailed protocol for a ligand-directed crosslinking method designed to identify the cell surface receptor of cell-binding antibodies and other protein ligands.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_10, © Springer Science+Business Media, LLC, part of Springer Nature 2019

Ligand-directed crosslinking methods, such as those using TRICEPS, HATRIC, or ASB [6–8,13], rely on a crosslinking event that is facilitated by binding of the ligand to the cell surface antigen, thus preferentially directing the crosslinks to the receptor of interest. Then a proteomic-style analysis is used to compare the crosslinked receptors for the ligand of interest compared to a control ligand, leading to the identification of putative cell surface receptors.

Here, we describe the synthesis and use of the ASB trifunctional crosslinker [8] that enables a ligand-directed crosslinking event, followed by subsequent transfer of a biotin from the crosslinker to the receptor of interest. The basic workflow for this method is depicted in Fig. 1.

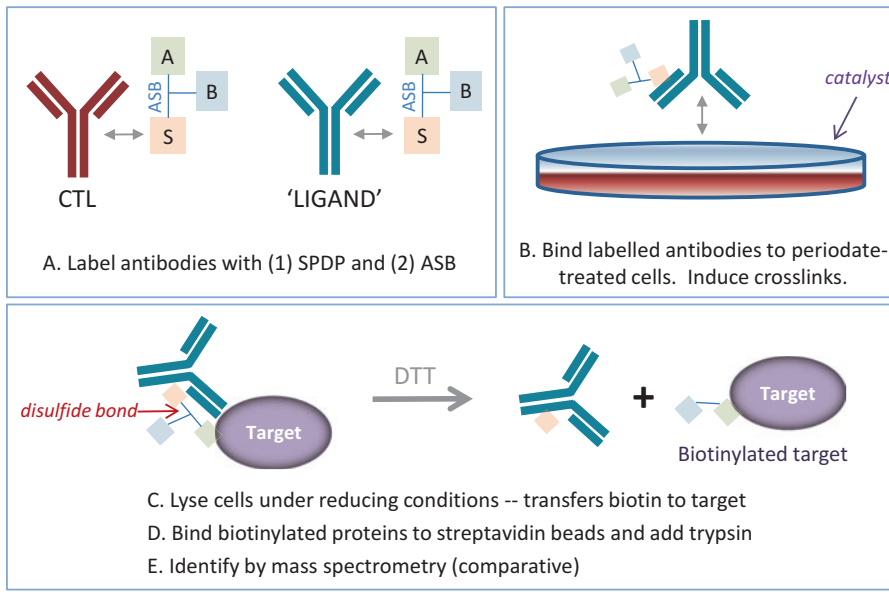


Fig. 1 ASB antigen identification workflow. (**a**) Two ligands are labelled with the ASB crosslinker using a two-step process. First, an SPDP-based crosslinker is added to the ligands through primary amines present in lysine residues on the ligand; then, ASB is added to the SPDP linker via reaction of its free sulfhydryl group with the dipyridyl group on the SPDP-derived linker. A control ligand is required to account for background binding and labelling of the cells, compared to the ligand of interest. (**b**) Next, the labelled ligands are allowed to bind to live cells that have been previously treated with periodate to produce aldehydes from the cell surface glycans. Crosslinks form between the aminooxy group on the ASB crosslinker and these aldehydes on the cell surface glycoproteins. (**c**) Cells are lysed under reducing conditions. This breaks the crosslinks and results in the transfer of the biotin to the crosslinked cell surface receptor. (**d**) Biotinylated proteins are purified using streptavidin and digested on-bead with trypsin. (**e**) A proteomic-style mass spectrometry analysis is used to identify cell surface proteins that were preferentially biotinylated in the ligand sample, relative to the control sample

2 Materials

Ensure that all materials are available before initiating experiment. Read all safety data sheets; handle and dispose of waste materials safely and appropriately.

2.1 ASB Crosslinker Synthesis

1. Fluorenylmethyloxycarbonyl-cysteamine-superacid sensitive resin (Fmoc-cysteamine-SASRIN).
2. Solid-phase peptide synthesis vessel with a Teflon valve and porous glass frit.
3. Fmoc-Asp(biotinyl-PEG)-OH.
4. Fmoc-NH-(PEG)2-COOH (20 atoms).
5. Dimethylformamide (DMF).
6. Dichloromethane (DCM).
7. Piperidine solution: Add 20% piperidine to DMF.
8. 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b] pyridinium 3-oxid hexafluorophosphate (HATU).
9. *N*,*N*-Diisopropylethylamine (DIPEA).
10. Boc-aminooxyacetic acid.
11. Aminooxyacetic acid.
12. 2,3,5-Collidine.
13. Cleavage solution: trifluoroacetic acid/water/triisopropylsilane (94:4:2, v/v/v).
14. Trifluoroacetic acid (TFA).
15. Cold ethyl ether: place 200 mL ethyl ether in a glass Erlenmeyer with glass stopper and place in a −20 °C freezer.
16. SepPAK columns: VacRC-C18—500 mg.
17. Properly vented vacuum dryer, suitable for organic solvents.
18. DNTB stock solution: 2 mM 5,5′-dithiobis(2-nitrobenzoate) and 50 mM sodium acetate in dd-H_2O; store at 4 °C.
19. Spectrometer capable of reading absorbance at 412 nm.
20. Freeze dryer.

2.2 Ligand Labelling with ASB Crosslinker

1. Ligands of interest (purified protein, 100 μg each).
2. Immobilized TCEP bead slurry.
3. Phosphate-buffered saline, pH 7.4.
4. EDTA stock solution: 20 mM ethylenediaminetetraacetic acid in water, adjusted to pH 8.
5. Forty percent acetonitrile (ACN): mix 40% acetonitrile in dd-H_2O (v/v).

6. Ninety-five percent ACN: mix 95% acetonitrile in dd-H_2O (v/v).
7. Empore C18 columns, 3 mL.
8. Amicon Ultra Centrifugal Filters: 30K cutoff for monoclonal antibodies. For smaller proteins a lower cutoff should be used, ideally 3× less than the protein molecular weight.
9. PEG_4-SPDP crosslinker: dissolved in DMSO at 8 mg/mL.
10. Glycine solution: 0.58 μg/μL glycine in PBS 7.4.
11. SpeedVac.

2.3 Cell Oxidation and Ligand Capture

1. Cells that ligand of interest can bind: ideally >10^8 cells should be available for each sample.
2. Live cell imaging solution (LCIS, Molecular Probes): LCIS is provided at pH 7.4.
3. Live cell imaging solution (LCIS), pH 6.5: adjust pH to 6.5 with hydrochloric acid (HCl).
4. Live cell imaging solution (LCIS), pH 8.0: adjust pH to 8.0 with sodium hydroxide (NaOH).
5. Oxidation solution: add 2 mM sodium meta-periodate to LCIS pH 6.5. Prepare fresh and protect from light.
6. *p*-Phenylenediamine stock solution: 1 M solution dissolved in DMSO.
7. Dry ice or liquid nitrogen.

2.4 Cell Lysis, Biotin Affinity Purification, and Trypsin Digestion

1. Lysis buffer: PBS 7.4 with 2% SDS, supplemented with 1:200 (v/v) protease inhibitor cocktail.
2. Probe sonicator.
3. Benzonase.
4. 50 mM ammonium bicarbonate (AMBIC), prepared fresh in water.
5. 200 mM dithiothreitol (DTT), prepared fresh in 50 mM AMBIC.
6. 375 mM iodoacetamide (IAA), prepared fresh in 50 mM AMBIC.
7. Streptavidin Plus UltraLink Resin (Pierce).
8. Mini-spin columns: empty with frit.
9. Vacuum manifold.
10. Trypsin: proteomics grade.

2.5 Peptide Identification and Relative Quantification by Mass Spectrometry

1. Liquid chromatograph with a C18 column, in line with a mass spectrometer.
2. Appropriate software to determine the relative quantification peptides between samples, such as MaxQuant [9] or PatternLab [10].
3. Appropriate software to identify peptides on the basis of their MS/MS fragmentation patterns.

3 Methods

3.1 Synthesis of ASB

ASB is synthesized using standard Fmoc-based solid-phase peptide synthesis methods, with some minor modifications. It may be possible to arrange for synthesis of this compound from one of the commercial vendors offering custom peptide synthesis services. Ensure that all safety precautions are taken and that work is performed in a certified chemical hood where appropriate. All reactions are at room temperature unless indicated otherwise.

1. Weigh out the equivalent of 0.3 mmol of Fmoc-cysteamine-SASRIN resin. Place resin in a peptide synthesis vessel.
2. Wash the resin (*see* **Note 1**) four times with DMF. Allow resin to swell in third wash by shaking for 15 min before removing solvent.
3. Deprotection: Add 7 mL of piperidine solution and shake gently for 5 min. Drain the solution and wash briefly with DMF. Add an additional 7 mL of piperidine solution and shake for 10 min. Repeat a third time. Wash the resin 3× with DMF, 3× with DCM, and 3× with DMF (*see* **Note 2**).
4. Coupling **step 1**:
   (a) Add to drained/semidry resin: 2.5 equivalents (0.75 mmol) of Fmoc-Asp(biotinyl-PEG)-OH and 2.5 eq. of HATU. Add DMF (5 mL) and mix. Add 3.5 eq. (1 mmol) of DIPEA and shake for 1 h.
   (b) Drain solution and add fresh Fmoc-Asp(biotinyl-PEG)-OH, HATU, DMF, and DIPEA as before. Incubate for an additional 45 min with shaking (*see* **Note 3**).
5. Wash: drain solution and wash the resin 6× with DMF.
6. Repeat deprotection (**step 3**).
7. Coupling **step 2**:
   (a) Add to drained/semidry resin: 2.5 eq. of Fmoc-NH-(PEG)2-COOH (20 atoms) and 2.5 eq. of HATU. Add DMF (5 mL) and mix. Add 3.5 eq. of DIPEA and shake for 1 h.
   (b) Drain solution, and add 1 eq. (0.3 mmol) of Fmoc-NH-(PEG)2-COOH (20 atoms) with 1 eq. HATU and 1.4 eq. DIPEA. Incubate for an additional 45 min with shaking.

8. Repeat washes (**step 5**).
9. Repeat deprotection (**step 3**).
10. Coupling **step 3**:
    (a) Add to drained/semidry resin: 2.5 eq. boc-aminooxyacetic acid and 2.5 eq. HATU. Add DMF (5 mL) and mix. Add 3.3 eq. 2,3,5-collidine. Incubate for 1 h.
    (b) Drain solution, and add fresh boc-aminooxyacetic acid, HATU, DMF, and collidine as in **step 10**a. Incubate for an additional hour.
11. Repeat washes (**step 5**).
12. Complete an additional five washes in DCM, and then dry the resin under vacuum (2–3 h to overnight).
13. Cleavage from resin: Add 10 eq. aminooxyacetic acid (3 mmol) to 5 mL cleavage solution, and add to the resin (*see* **Note 4**). Shake for 2 h at room temperature.
14. Ether precipitation:
    (a) Drip the cleavage solution from **step 13** into 50 mL Falcon tubes containing 40 mL of cold ethyl ether.
    (b) Add 2 mL of cleavage solution to the resin, followed by 2 mL of neat TFA, and allow to drip into another 50 mL Falcon tum containing 40 mL of cold ethyl ether.
    (c) Place ether-filled tubes at −20 °C for 10–20 min to aid precipitation. Centrifuge at 1100 × *g* for 20 min and decant ether.
    (d) Resuspend pellet in 5 mL of neat TFA. ASB may require up to 1 h of shaking to fully resuspend.
    (e) Combine fractions from the two tubes, add an additional 40 mL of cold ether, and precipitate again. After centrifuging, decant the supernatant and dry pellet under nitrogen.
    (f) Dissolve the product in 8 mL of distilled water, and aliquot 2 mL into each of four vials.
    (g) Freeze on dry ice and lyophilize. Store at −20 °C.
15. SepPak purification (*see* **Note 5**):
    (a) Resuspend one lyophilized aliquot of ASB into 12 mL of ddH_2O.
    (b) Equilibrate four SepPak columns with 10 mL of ddH_2O, followed by 10 mL of 95% ACN/ddH_2O and 20 mL of ddH_2O.
    (c) Slowly load 3 mL of the resuspended ASB onto each column.
    (d) Wash the columns four times with 3 mL of ddH_2O.

(e) For elution, add 3 mL of 20% acetonitrile to each column, and collect the flow-through in a clean tube containing 350 mg of free aminooxyacetic acid, combining elution fractions from all columns.

16. Run an Ellman assay to quantify the ASB based on its free sulfhydryl, using L-cysteine to make a standard curve [11].
    (a) Prepare working reagent by adding 300 µL of DNTB stock solution and 600 µL of 1 M Tris–HCl pH 8 to 5 mL of dd-H_2O.
    (b) Prepare a 0.33 µg/µL solution of L-cysteine in dd-H_2O, fresh.
    (c) Dry down a 5 and 12 µL aliquot of the 20% flow-through SepPak fractions, and resuspend in 15 µL of dd-H_2O.
    (d) Prepare a standard curve with seven samples each containing 0, 0.5, 1, 2, 3, 4, and 5 µg of L-cysteine in 15 µL of dd-H_2O.
    (e) To each ASB and standard curve sample, add 200 µL of working reagent, and incubate at room temperature for 5 min.
    (f) Read absorbance at 412 nm, and calculate amount of ASB in each fraction based on the free sulfhydryl content (nmol/µL), using the known L-cysteine samples as standards.
17. Aliquot 500 µg of ASB into glass vials and lyophilize. Store at −20 °C until use.

3.2 Ligand Labelling with ASB Crosslinker

These instructions are written assuming that two monoclonal antibodies are going to be labelled. We suggest labelling antibodies fresh before each use. The labelling steps can easily be adapted to other types of ligands—however, the amount of linker and ASB may need to be adapted depending on the protein size and sequence. Mass spectrometry on intact labelled proteins can be helpful for determining optimal ratios for each protein.

1. Prepare ASB crosslinker to reduce any disulfide bonds that have formed, and remove the free aminooxyacetic acid.
    (a) Place 0.5 mL of immobilized TCEP bead slurry in a 1.5 mL tube, and wash 2× with PBS pH 7.4 + 1 mM EDTA. Leave beads as dry as possible.
    (b) Dissolve an aliquot of dried ASB in PBS 7.4 + 1 mM EDTA to produce a final concentration of 1 mg/mL ASB.
    (c) Add 250 µg of the diluted ASB to the immobilized TCEP, and incubate with rotation at room temperature for 1 h. When ready to proceed, remove the reduced ASB solution from the beads by running through an empty mini-spin

column. Any remaining ASB can be frozen and stored at −20 °C for later use.

(d) Clean ASB through an Empore C18 column, as follows (*see* **Note 6**):

- Condition column: run through 250 µL of 95% ACN/water.
- Equilibrate column: run through 2 × 500 µL water.
- Add 250 µL of ASB sample to each column. If more ASB is required, we recommend using two Empore columns to avoid overloading.
- Wash column: run through 3 × 500 µL water.
- Elute ASB: add 250 µL 40% ACN, and collect flow-through in a clean tube by centrifugation at 200 × *g* for 3 min.
- Redistribute elution volume into clean tubes: 60 µL (~60 µg) for each 100 µg of mAb to label.
- SpeedVac 60 µg aliquots until the volume in each tube is 15–20 µL.

2. Label 100 µg of each antibody with the PEG_4-SPDP linker (*see* **Note 7**).

(a) Take 100 µg of each protein, and bring volume up to 350 µL with PBS pH 7.4.

(b) Buffer exchange each protein into PBS pH 7.4 using Amicon filters, according to the manufacturer's instructions. Transfer protein from top of filters into clean tubes, and add PBS 7.4 to a final volume of 50 µL. Store labelled, used Amicon filters at 4 °C after adding 500 µL PBS to the top of the filter. These will be used in **step 4**.

(c) Add 1.4 µL of PEG_4-SPDP to each protein, and incubate for 30 min at room temperature.

(d) Quench with 3.0 µL of glycine solution, and incubate 30 min more at room temperature.

(e) Add 1 mM EDTA (final concentration) by adding 2.75 µL of 20 mM EDTA pH 8.

(f) Add each labelled protein to one 60 µg aliquot of the prepared ASB, and incubate O/N at 4 °C.

3. Remove excess ASB by buffer exchanging into PBS 7.4 reusing the Amicon filters saved above. Transfer labelled protein from top of filter to a new tube, and adjust the volume of each labelled protein to 100 µL with PBS 7.4.

3.3 Cell Oxidation and Ligand Capture

This section describes the oxidation of glycans on the surface of live cells, resulting in the production of aldehydes, and then the

subsequent incubation of the labelled ligands with the cells to induce crosslink formation. In theory, more crosslinks will be formed with proteins that are in close proximity to the antibody/ligand of interest upon cell binding.

1. Collect tissue culture cells in LCIS buffer (*see* **Note 8**). Once collected, place cells on ice. Pellet cells by centrifugation at 4 °C (*see* **Note 9**), and resuspend in 40 mL of cold LCIS pH 6.5 to wash away most unbound proteins.
2. Pellet cells, and resuspend in 50 mL of oxidation solution. Oxidize cells at 4 °C in the dark for 15 min under constant gentle agitation (*see* **Note 10**).
3. Pellet cells, and resuspend in 50 mL of LCIS pH 7.4. Repeat this step with resuspension in LCIS pH 8.0 (*see* **Note 11**).
4. Pellet cells and resuspend in 20 mL of LCIS pH 8.0. Split cell suspension into 2 × 10 mL in 15 mL centrifuge tubes, being careful to keep cells fully and evenly suspended. Splitting the cells after oxidation ensures that both cell populations are as similar as possible for the binding step.
5. Add each ASB-labelled monoclonal antibody to one tube containing resuspended cells.
6. Incubate ligands with cells for 1.5 h at 4 °C with gentle rocking or rotation.
7. Add 10 mM p-phenylenediamine, and incubate at 4 °C for an additional 25 min (*see* **Note 12**).
8. Pellet cells to remove most of the unbound ligand. Resuspend the cells in 10 mL of cold LCIS pH 7.4, and repeat this step two additional times.
9. Flash freeze the semidry cell pellet with liquid nitrogen or in a dry ice/ethanol bath. Store cells at −80 °C indefinitely.

3.4 Cell Lysis, Biotin Affinity Purification, and Trypsin Digestion

This section describes the isolation of biotinylated proteins from the labelled cells, followed by a trypsin digestion step. At the end of this section, the samples will be ready for analysis by a proteomic mass spectrometry method.

1. Resuspend each cell pellet in a volume of lysis buffer that is approximately 10× the volume of the cell pellet (the "cell pellet volume"). Homogenize the lysate by passing it through a 22G3/4 needle five times. Incubate lysate at room temperature for 30 min with gentle agitation, taking care to minimize foaming.
2. Further dilute sample by adding 1/3 cell pellet volume of PBS pH 7.4 supplemented with 1:200 protease inhibitor cocktail. Sonicate for approximately 30 s (*see* **Note 13**).

3. Add 1/1000 volume of Benzonase to each tube, and incubate at room temperature for another 45 min.
4. Centrifuge at 4500 × *g* and transfer supernatant in new 15 mL tube.
5. Add 1/20 volume of 200 mM DTT for a final concentration of 10 mM. Heat for 10 min in boiling water, and then let it cool to room temperature for 10 min. This will reduce all disulfide bonds resulting in the cleavage of any crosslinks with the biotin left on the ligand side.
6. Add 1/10 volume of 375 mM IAA to lysate, and incubate at room temperature in dark 25 min with gentle agitation.
7. Add 1/20 volume of 200 mM DTT (final concentration of 20 mM) to quench alkylating reagent.
8. Dilute lysate 5.3× with 50 mM AMBIC in 50 mL tubes.
9. For the biotin affinity purification, transfer 200 µL of Streptavidin Plus UltraLink Resin slurry to a new tube, and wash the resin three times with 50 mM AMBIC.
10. Aliquot ~20 µL of settled streptavidin resin into each diluted cell lysate sample, and incubate for 1 h at room temperature with gentle rotation.
11. Carefully transfer the streptavidin resin to an empty mini-spin column, working to minimize bead loss. Using a vacuum manifold will facilitate this process.
12. Wash the streptavidin resin as follows (*see* **Note 6**):
    (a) 5× with 500 µL of PBS + 0.5% SDS.
    (b) 5× with 500 µL of 2 M urea in 50 mM AMBIC.
    (c) 10× with 500 µL of 50 mM AMBIC.
13. Transfer the resin back into an Eppendorf tube, and carefully remove the supernatant.
14. Add 80 µL of 50 mM AMBIC containing 1.3 µg of trypsin to the resin for each sample, and incubate O/N at 37 °C.
15. Remove resin, and collect supernatant containing the digested peptides in a new tube (*see* **Note 14**). Store digest at −80 °C until analysis by mass spectrometry.

3.5 Peptide Identification and Relative Quantification by Mass Spectrometry

Any reliable proteomics-style method for relative quantification of peptides between two samples would be suitable for this analysis. In this section, we focus on points for consideration that are specific to analysis of ASB biotin transfer data rather than providing a specific detailed method. Many mass spectrometry centers at academic institutions offer proteomic services, including analysis and advice.

1. Peptides should be separated by a sensitive LC system. Typically, C18 chemistries are used to separate peptides, and

Table 1
Tryptic streptavidin peptides

m/z	z	Peptide
402.544	3	STLVGHDTFTK
491.482	4	NAHSATTWSGQYVGGAEAR
603.312	2	STLVGHDTFTK
654.973	3	NAHSATTWSGQYVGGAEAR
678.681	3	INTQWLLTSGTTEANAWK
718.676	3	YDSAPATDGSGTALGWTVAWK
981.956	2	NAHSATTWSGQYVGGAEAR
1017.518	2	INTQWLLTSGTTEANAWK
1077.511	2	YDSAPATDGSGTALGWTVAWK

These peptides are often found to be quite abundant in samples using an on-bead trypsin digest. Addition of these peptides to an exclude list during LC-MS analysis can improve protein identifications

the LC system is placed directly in line with a mass spectrometer. The peptide mixtures from an ASB transfer experiment contain large amounts of streptavidin peptides derived from the purification step. We recommend utilizing a system with large dynamic range to allow detection of smaller peptide signals in among the larger background signal. We also suggest injecting samples at least two times. In the second injections, an exclude list for high abundance streptavidin peptides can be utilized. The peptides we regularly exclude on an LTQ-Orbitrap system are listed in Table 1.

2. To identify peptides of interest, the relative intensity of peptides in each of the two ASB samples must be compared. We suggest using label-free quantification based on the MS signal intensity in a high-resolution mass spectrometer, but other methods such as spectral counting may be suitable. With ligands that bind tightly to abundant targets on cells, exploration of the Mascot results may be sufficient for comparison. However, most ligands we have tested require a closer look at the comparative MS data. We have found that MSight [12] (https://web.expasy.orb/MSight) can be a valuable and accessible tool for comparison of MS profiles between samples.
3. Peptide identifications are typically obtained through data-dependent MS/MS with dynamic exclusion. We have also found that include list methods, focused on ions of interest, such as those identified as differentially expressed through MSight analysis, can be valuable.
4. Some peptides that are differentially expressed between the two samples may belong to the ligand itself, so the ligand

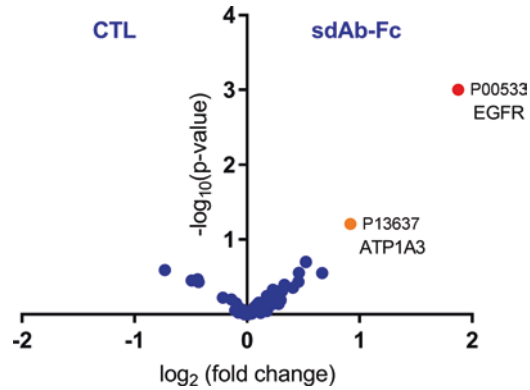


Fig. 2 Volcano plot visualization of ASB antigen identification results at the protein level. In this experiment, the ligand of interest was a single-domain antibody fused to the Fc region of an IgG (sdAb-Fc), and the control (CTL) was a similar protein construct that did not bind to Panc-1 cells. Identified proteins were filtered to remove keratins and contaminating non-cell surface housekeeping proteins before statistical analysis to identify differentially expressed proteins. The results show that the protein with the strongest differential expression in the sdAb-Fc sample is EGFR, represented by the red dot. Another protein, ATP1A3, is also differentially expressed to a lower degree and may represent a secondary hit or a "near neighbor" of EGFR

sequence should be included in any database search methods used. It may also be helpful to digest some ASB-labelled ligand with trypsin for comparison with the ASB results and elimination of these signals as irrelevant background.

5. Full biological replicates can greatly improve statistics and allow identification of weaker peptide signals that are differently expressed in the LIGAND and CTL samples. As depicted in Fig. 2, a volcano plot with the $\log_2$ (fold change) on the *x*-axis and the $-\log_{10}(p\text{-value})$ on the *y*-axis can be a nice visualization tool for the antigen identification results. In this example, two proteins (EGFR and ATP1A3) clearly stand out as differentially expressed in the sdAb-Fc sample, as highlighted by their position at the upper-outer quadrant of the volcano plot.

4 Notes

1. For each wash step, solvent equivalent to approximately 3× bead height (5–7 mL at the scale discussed here) is added to the resin and mixed gently, and the solvent is drained into an appropriate waste container. This step is repeated as required. If time permits, it is often helpful to allow for the penultimate wash to incubate for 5 min prior to draining to ensure more thorough removal of unbound materials.

2. We found that the beads were rather "sticky" at this stage. At each wash step, it may be helpful to add the wash solvent such that it cascades down the inside of the reactor walls onto any beads stuck on the wall of the reactor. We found that the DCM wash steps helped to alleviate some of the stickiness.
3. We found that the two-stage coupling reaction improved yields. While most synthesis protocols use a Kaiser test to determine whether coupling is complete, we found that this test was unreliable using these larger PEG amino acids.
4. Addition of an excess of free aminooxyacetic acid acted as a quencher for environmental aldehydes which otherwise reacted with the aminooxy group on the ASB linker.
5. This is a simple method for crude purification that does not require specialized equipment. Alternatively, a chromatography purification can be performed here. We have found that inclusion of free aminooxy groups, such as those found in aminooxyacetic acid (Aoa), may help to protect the aminooxy group of ASB from scavenging environmental aldehydes. Avoiding low pH conditions also helps protect the aminooxy group, which is why we do our purification without the addition of acid. In fact, in the absence of low pH conditions, addition of Aoa may not be entirely necessary. However, careful monitoring of the ASB to ensure no contamination from environmental aldehydes is recommended in this case. If ASB is stored without free Aoa, the "clean" step in each ligand labelling experiment can be avoided.
6. The use of a vacuum manifold can expedite the wash steps.
7. Other amine-pyridyldithio crosslinkers can be used in this step. We have found the PEG_4 version of SPDP to be less likely to induce aggregation in the labelled protein than LC-SPDP. Longer or shorter linker lengths can be tested, if desired for your application.
8. We recommend using Accutase to remove adherent cells from plates. Scraping gently into PBS can also be suitable for some cell lines. We have found improved cell viability using Live Cell Imaging Solution during this step, but PBS can also be used.
9. For most cell lines, cells can be pelleted by centrifugation at $200 \times g$ for 5 min at 4 °C, and the supernatant can be safely decanted or aspirated.
10. Cell aggregation may be visible at this step, and a portion of the cells may be floating after the next centrifugation step.
11. The pH modification is designed to slow down the uncatalyzed reaction between the aminooxy group and oxidized cell surface glycans while also allowing ligand binding at a physiological pH.

12. Cells may become slightly purple/gray colored.
13. The level of sonication may vary depending on the sonicator. Choose a setting and pulse sequence that does not heat the sample and that does not foam excessively.
14. Beads can be efficiently and completely separated from the digest liquid by using an empty spin column that has been pre-wetted in AMBIC buffer.

Acknowledgments

We would like to thank the National Research Council Canada for supporting this research.

References

1. Zhao A, Tohidkia MR, Siegel DL, Coukos G, Omidi Y (2016) Phage antibody display libraries: a powerful antibody discovery platform for immunotherapy. Crit Rev Biotechnol 36(2): 276–289. https://doi.org/10.3109/07388851.2014.958978
2. Roodink I, Franssen M, Zuidscherwoude M, Verrijp K, van der Donk T, Raats J, Leenders WP (2010) Isolation of targeting nanobodies against co-opted tumor vasculature. Lab Investig 90(1):61–67. https://doi.org/10.1038/labinvest.2009.107
3. Stutz CC, Zhang X, Shusta EV (2014) Combinatorial approaches for the identification of brain drug delivery targets. Curr Pharm Des 20(10):1564–1576
4. Xiao X, Chen Y, Mugabe S, Gao C, Tkaczyk C, Mazor Y, Pavlik P, Wu H, Dall'Acqua W, Chowdhury PS (2015) A novel dual expression platform for high throughput functional screening of phage libraries in product like format. PLoS One 10(10):e0140691. https://doi.org/10.1371/journal.pone.0140691
5. Ten Haaf A, Pscherer S, Fries K, Barth S, Gattenlohner S, Tur MK (2015) Phage display-based on-slide selection of tumor-specific antibodies on formalin-fixed paraffin-embedded human tissue biopsies. Immunol Lett 166(2):65–78. https://doi.org/10.1016/j.imlet.2015.05.013
6. Frei AP, Jeon OY, Kilcher S, Moest H, Henning LM, Jost C, Pluckthun A, Mercer J, Aebersold R, Carreira EM, Wollscheid B (2012) Direct identification of ligand-receptor interactions on living cells and tissues. Nat Biotechnol 30(10):997–1001. https://doi.org/10.1038/nbt.2354
7. Frei AP, Moest H, Novy K, Wollscheid B (2013) Ligand-based receptor identification on living cells and tissues using TRICEPS. Nat Protoc 8(7):1321–1336. https://doi.org/10.1038/nprot.2013.072
8. Tremblay TL, Hill JJ (2017) Biotin-transfer from a trifunctional crosslinker for identification of cell surface receptors of soluble protein ligands. Sci Rep 7:46574. https://doi.org/10.1038/srep46574
9. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11(12):2301–2319. https://doi.org/10.1038/nprot.2016.136
10. Carvalho PC, Lima DB, Leprevost FV, Santos MD, Fischer JS, Aquino PF, Moresco JJ, Yates JR 3rd, Barbosa VC (2016) Integrated analysis of shotgun proteomic data with PatternLab for proteomics 4.0. Nat Protoc 11(1):102–117. https://doi.org/10.1038/nprot.2015.133
11. Ellman GL (1959) Tissue sulfhydryl groups. Arch Biochem Biophys 82(1):70–77
12. Palagi PM, Walther D, Quadroni M, Catherinet S, Burgess J, Zimmermann-Ivol CG, Sanchez JC, Binz PA, Hochstrasser DF, Appel RD (2005) MSight: an image analysis software for liquid chromatography-mass spectrometry. Proteomics 5(9):2381–2384. https://doi.org/10.1002/pmic.200401244
13. Sobotzki N, Schafroth MA, Rudnicka A, Koetemann A, Marty F, Goetze S, Yamauchi Y, Carreira EM, Wollscheid B (2018) HATRIC-based identification of receptors for orphan ligands. Nat Commun 9 (1): 1519. https://doi.org/10.1038/s41467-018-03936-z



Chapter 11

Whole-Genome Phage Display Libraries: A Powerful Tool for Antigen Discovery

Elisa Beghetto and Nicola Gargano

Abstract

In the last two decades, phage display technology has been used for investigating complex biological processes and isolating molecules of practical value in several applications. Bacteriophage lambda, representing a classical cloning and expression system, has also been exploited for generating display libraries of small peptides and protein domains. More recently, large cDNA and whole-genome lambda display libraries of human pathogens have been generated for the discovery of new antigens for biomedical applications. Here, we describe the construction of a whole-genome library of a common pathogen—*Streptococcus pneumoniae*—and the use of this library for the molecular dissection of the human B-cell response against bacterial infection and colonization.

Key words Bacteriophage lambda, Phage display library, Genomic DNA display, *Streptococcus pneumoniae*, Antigen discovery

1 Introduction

Lambda bacteriophage has been demonstrated to be the system of choice to display complex complementary deoxyribonucleic acid (cDNA) libraries. Libraries with a complexity of 10^7–10^8 independent clones can be easily constructed using efficient in vitro packaging systems.

The head decoration protein gpD (11.4 kDa, 405–420 copies per capsid) represents the ideal fusion partner, as it has been demonstrated that gpD can tolerate both amino- and carboxyl-terminal insertions of peptides and protein domains accessible for ligand interaction without interfering in phage replication and assembly of infective bacteriophages [1–4].

One of the most commonly used vectors for displaying large genomic and cDNA repertoires in lambda bacteriophage, named λKM4, is based on a "two-gene system," which is characterized by the presence of two copies of the D gene: a genomic copy harboring an amber mutation (a stop codon causing the premature

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_11, © Springer Science+Business Media, LLC, part of Springer Nature 2019

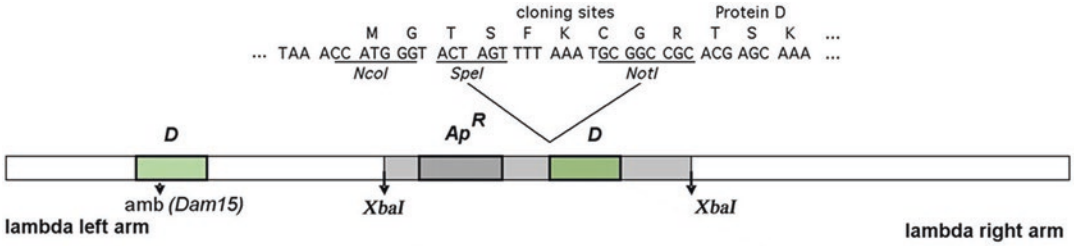


Fig. 1 Genetic map of λKM4 vector with sequence of the cloning sites. *ApR* ampicillin resistance, *amb* amber mutation

termination of gpD synthesis, when the gpD expression is performed in an "non-suppressor" bacterial strain) at the 5′ end, followed by an additional copy, controlled by an inducible promoter, containing a cloning cassette at the 5′ end (*see* Fig. 1) [5, 6]. Accordingly, bacteriophages grown on a suppressor bacterial strain (translating the amber mutation in a phenylalanine residue) display an array of wild-type gpD (encoded by the genomic copy) and recombinant gpD fused to foreign polypeptides (encoded by the additional copy). Such an array of wild-type proteins and recombinant fusion products is necessary to avoid the full display of recombinant proteins (when all of the capsidic gpD copies are fused to the exogenous polypeptide), which can cause steric hindrance and disassembly of the viral capsid. The vector also harbors an antibiotic resistance gene (i.e., β-lactamase), allowing the growth of recombinant clones as antibiotic-resistant lysogenic colonies.

Most of the selection schemes for screening lambda libraries utilize protocols originally developed for filamentous phage applications [7, 8]. Briefly, the target molecules linked to solid-phase matrices (i.e., polypropylene plates or tubes, sepharose matrices, or magnetic beads) are challenged with a suspension of phage particles representing the library. After incubating the mixture for several hours, phage-ligand complexes are used to infect freshly cultured bacteria, and the phage progeny released by host cells is directly screened for isolating individual clones or further processed for another round of affinity selection. In order to eliminate most of the low-affinity interactors and to enrich phage population for specific binders, the selection procedure generally includes 2–3 rounds of affinity selection.

In the last two decades, the lambda display approach has been successfully employed to identify and characterize antigens involved in the human immune response to infections caused by several pathogens, including protozoa, viruses, and bacteria, and to identify tumor-associated antigens by screening human cDNA libraries with sera from tumor patients [5, 6, 9–16]. This was achieved by selecting the antigenic regions harboring B-cell epitopes via a direct challenge of protein-fragment display libraries with the

whole antibody repertoire of individuals affected by the corresponding disease (i.e., antibodies present in plasma or serum).

Recently, the efficiency of lambda display has been combined with the power of next-generation sequencing into a new platform, named PROFILER (standing for "Phage-based Representation OF ImmunoLigand Epitope Repertoire"), which allows the characterization of antibody repertoires in polyclonal mixtures such as serum samples from vaccinated individuals [17]. In a proof of concept study, the use of this technology allowed the identification of the epitope recognized by a monoclonal antibody against *Neisseria meningitidis* and provided, in addition, clues for the functional characterization of the epitope [18].

Finally, the potential of lambda display as oral vaccine delivery vehicle for peptide antigens has been recently explored [19]. A recombinant lambda phage was constructed to display three immunogenic epitopes from the cervid prion protein on the phage capsid gpD. Targeted delivery of purified recombinant phages to intestinal segments of mice induced IgA responses to all three peptide epitopes. Further, delivery of bacteria expressing soluble epitope gpD fusion protein also induced epitope-specific IgA responses in the small intestine.

In this chapter, the experimental design for investigating the human antibody response to *Streptococcus pneumoniae* (pneumococcus), a ubiquitous gram-positive bacterium causing invasive diseases such as pneumonia, sepsis, and meningitis [20–22], will be described. In particular, the construction of a *Streptococcus pneumoniae* lambda display library of genomic DNA (gDNA) fragments and the challenge of the library with sera from patients hospitalized for *S. pneumoniae* diseases will be detailed (*see* Fig. 2), followed by a preliminary molecular characterization of the selected recombinant clones, which represent B-cell antigens or antigenic regions containing B-cell epitopes recognized by antibodies of infected individuals.

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of 18 MΩ cm at 25 °C) and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing waste materials. Use disposable filter tips, especially handling phage suspensions.

2.1 Bacterial Strains, Vectors, and Human Sera

1. *Escherichia coli* BB4 and LE392 strains are commercially available (i.e., Invitrogen). Any bacterial strain susceptible to lambda bacteriophage infection is suitable for the display library construction.

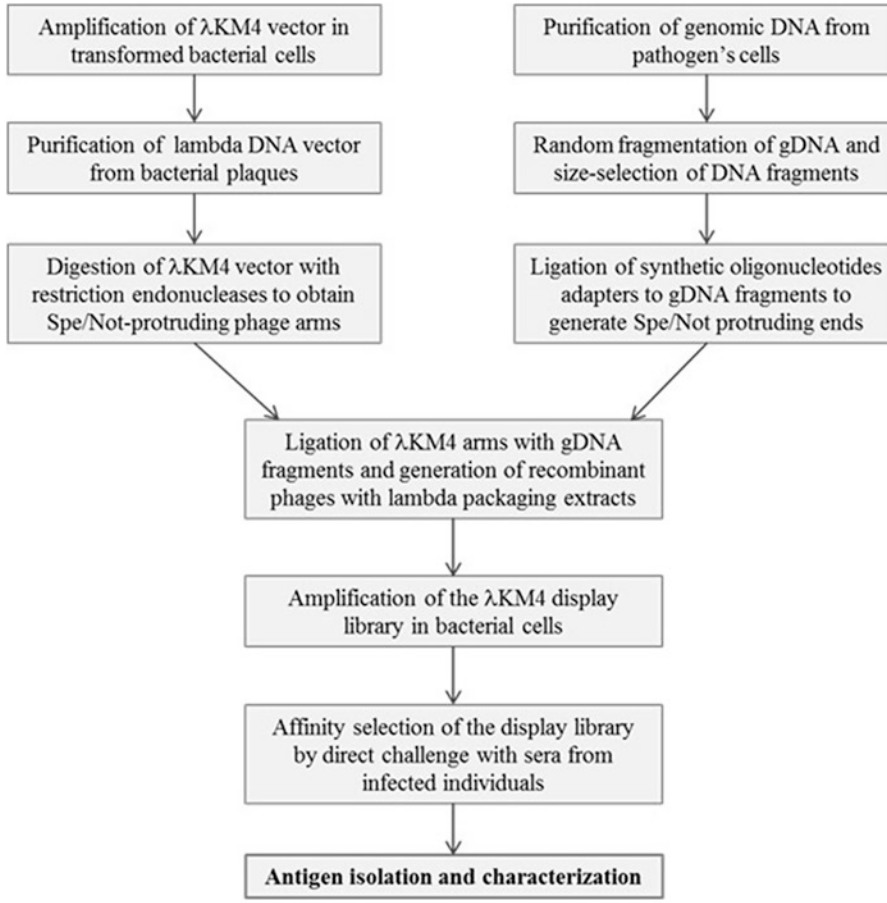


Fig. 2 Flowchart of library construction and antigen selection

2. λKM4 vector is not commercially available, but it can be provided by the authors of the present article for non-commercial purposes. Alternatively, it could be provided by other authors [6].
3. Human sera or plasma from *S. pneumoniae*-infected individuals is collected from hospitals, blood banks, etc. (for the data presented here, sera have been provided by the Clinic of Infectious Diseases of Siena University, Italy).

2.2 Buffers and Reagents

1. Phosphate-buffered saline (PBS).
2. 3 M Sodium acetate (CH_3COONa) stock solution: dissolve 408.3 g of CH_3COONa in 1 L of H_2O. Adjust pH to 5.2 by adding acetic acid. Autoclave and store the stock solution at room temperature (up to 6 months).
3. Isopropyl alcohol.
4. Ethyl alcohol.

5. 0.5 M Ethylenediaminetetraacetic acid (EDTA) stock solution: dissolve 186 g of EDTA in 1 L of distilled, deionized H_2O. Adjust pH to 8.0 by using NaOH.
6. Adenosine-5′-triphosphate (ATP).
7. Bovine serum albumin (BSA).
8. Deoxyribonucleotide triphosphates (dNTPs).
9. Agarose gel electrophoresis: all reagents/devices (1.2% precast agarose gels, electrophoresis buffer, running device) for standard gel electrophoresis are commercially available (e.g., purchased from Sigma-Aldrich).
10. SM buffer: dissolve 5.8 g of NaCl, 2.0 g of $MgSO_4[7H_2O]$, 50 mL of 1 M Tris–HCl (pH 7.5), and 5.0 mL of 2% gelatine with H_2O to a final volume of 1 L. Autoclave and store the solution at room temperature for a maximum of 6 months.
11. 2% Maltose stock solution: dissolve 2 g of maltose in 100 mL of H_2O.
12. 1 M Magnesium sulfate ($MgSO_4$) stock solution: dissolve 246.5 g of $MgSO_4$ in 1 L of H_2O. Store at room temperature for up to 12 months.
13. 5 M Sodium chloride (NaCl) stock solution: dissolve 292.2 g of NaCl in 1 L of H_2O.
14. Polyethylene glycol 8000 (PEG-8000; Sigma-Aldrich).
15. Dimethyl sulfoxide (DMSO).
16. Chloroform.

2.3 Bacterial Media

1. Luria Broth (LB; all regents purchased from Sigma-Aldrich): dissolve 10 g of NaCl, 10 g of bacto tryptone, and 5 g bacto yeast extract in 1 L of H_2O. Adjust to pH 7.0 with 5 N NaOH. Autoclave and store the medium at room temperature for a maximum of 4 months.
2. NZY Broth (all reagents from Sigma-Aldrich): dissolve 5 g of NaCl, 2 g of $MgSO_4[7H_2O]$, 5 g of yeast extract, and 10 g of casein hydrolysate in H_2O (final volume of 1 L). Adjust the pH to 7.5 with NaOH. Autoclave and store the medium at room temperature for a maximum of 4 months.
3. NZY Top Agar and NZY Agar: NZY Broth supplemented with 0.7% (w/v) or 1.5% (w/v) of agar (Sigma-Aldrich), respectively.
4. Tryptic Soy Broth (TSB; Becton Dickinson).

2.4 Synthetic Oligonucleotides

The following oligonucleotides can be purchased from commercial suppliers (i.e., Sigma-Aldrich custom orders) and stored in H_2O at −20 °C in small aliquots until use. It should be noted that such oligonucleotides are specific for cloning DNA inserts into λKM4

vector [9], independently from the source (i.e., gDNA or cDNA) and the pathogen of interest (i.e., *S. pneumoniae*).

K185 5′-CTAGTCGTGCTGGCCAGC-3′.

K186 5′-GCTGGCCAGCACGA-3′.

K187 5′-CTAGTCGTGCTGGCCAGCT-3′.

K188 5′-AGCTGGCCAGCACGA-3′.

K189 5′-CTAGTCGTGCTGGCCAGCTG-3′.

K190 5′-CAGCTGGCCAGCACGA-3′.

K191 5′-TCTGGTGGCGGTAGC-3′.

K192 5′-GGCCGCTACCGCCACCAGA-3′.

K193 5′-TTCTGGTGGCGGTAGC-3′.

K194 5′-GGCCGCTACCGCCACCAGAA-3′.

K195 5′-TTTCTGGTGGCGGTAGC-3′.

K196 5′-GGCCGCTACCGCCACCAGAAA-3′.

λfor: 5′-GGGCACTCGACCGGAATTATCG-3′.

λrev: 5′-CTCTCATCCGCCAAAACA GCC-3′.

2.5 Laboratory Instruments

1. Spectrophotometer, allowing the measurement in both ultraviolet (UV) and visible wavelength spectra, is available from many suppliers (i.e., Eppendorf).
2. UV transilluminator (Vilber Lourmat; LabTech).
3. Automated microplate ELISA reader having absorbance detection modules/filters for 450 and 620 nm wavelengths (Bio-Rad).
4. Automated ELISA washer (Tecan).
5. Power supply and device for agarose gel electrophoresis (all apparatus for standard electrophoresis can be purchased from Sigma-Aldrich).
6. Other equipment for molecular biology laboratory: incubators (i.e., cabinet and shaking incubators for growing bacteria in plates and flasks, respectively), centrifuges (for tubes and bottles), thermomixer (Eppendorf), pipettes (Gilson), etc.

2.6 DNA Purification Kits and Lambda Packaging Kits

1. Large-scale preparation of bacteriophage λKM4 vector can be obtained by using commercially available kits (i.e., Lambda DNA Purification Kit purchased from Agilent Technologies).
2. Wizard SV Gel and polymerase chain reaction (PCR) cleanup system (provided by Promega) for purification of PCR products and other enzyme-processed DNA fragments.
3. The packaging of lambda bacteriophages for the construction of display libraries can be performed by using commercially

available packaging kits (i.e., MaxPlax™ Lambda Packaging Extracts purchased from Lucigen).

2.7 Enzymes and Buffers

All of the following enzymes, as well as the corresponding reaction buffers, are purchased from New England Biolabs:

1. SpeI restriction endonuclease.
2. NotI restriction endonuclease.
3. Deoxyribonuclease-(ribonuclease-free) (DNaseI-(RNase-free)).
4. T4 DNA ligase and corresponding 10× buffer.
5. T4 Polynucleotide Kinase and corresponding 10× buffer.
6. T4 DNA polymerase and corresponding 10× buffer.

2.8 Biopanning Reagents and Buffers

1. Dynabeads Protein G (Thermo Fisher Scientific).
2. Fetal Calf Serum (FCS; Thermo Fisher Scientific).
3. Biopanning coating buffer: 0.1 M NaH_2PO_4/Na_2HPO_4, pH 8.
4. Biopanning washing buffer: 1% (v/v) Triton X-100 and 10 mM $MgSO_4$ in PBS.
5. Biopanning blocking buffer: 5% (w/v) nonfat dry milk in PBS, 0.25% Triton X-100, and 10 mM $MgSO_4$.

2.9 Phage Enzyme-Linked Immunosorbent Assay (ELISA) Reagents, Antibodies, and Buffers

1. MaxiSorp multiwell plates (Thermo Fisher Scientific).
2. Polyclonal anti-lambda antibodies (developed in rabbits) (*see* **Note 1**).
3. Secondary antibody: goat antihuman IgG horseradish peroxidase (HRP)-conjugated antibodies (Sigma-Aldrich).
4. Chromogenic substrate: tetramethylbenzidine (TMB) liquid substrate system (Sigma-Aldrich).
5. Phage ELISA coating buffer: 50 mM $NaHCO_3$ pH 9.6.
6. Phage ELISA washing buffer: 0.05% (v/v) Tween 20 in PBS.
7. Phage ELISA blocking buffer: 5% (w/v) nonfat dry milk and 0.05% (v/v) Tween 20 in PBS.
8. Stop solution: 2 M H_2SO_4.

2.10 Immunoscreening Reagents, Antibodies, and Buffers

1. Nitrocellulose filters (Schleicher & Schuell).
2. Secondary antibody: goat antihuman IgG alkaline phosphatase (AP)-conjugated antibodies (Sigma-Aldrich).
3. Chromogenic substrates: nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) substrates (Sigma-Aldrich).
4. Immunoscreening washing buffer: 0.05% Tween 20 in PBS.
5. Immunoscreening blocking buffer: 5% nonfat dry milk and 0.1% Tween 20 in PBS.

3 Methods

3.1 Preparation of Lambda Display Transfer Vector (λKM4)

1. Organize a medium-/large-scale preparation of λKM4 vector (i.e., 50–100 μg of DNA) for cloning the exogenous DNA fragments into bacteriophage genome and generating high-complexity display libraries; grow the bacteriophage in *E. coli* BB4 cells cultured in top agar medium, thus allowing the formation of phage plaques onto the lawn of bacterial plates.
2. Purify λKM4 DNA from phage plaques using commercial kits (i.e., Lambda DNA Purification Kit from Agilent Technologies); strictly adhere to manufacturer's instruction for growing conditions (initial amounts of BB4 cells and lambda bacteriophage, temperature and length of incubation, etc.).
3. Check the integrity of λKM4 genome by running 1 μg of the purified DNA with agarose gel electrophoresis (agarose gel, 1.2% (w/v)).
4. Incubate for 2 h at 37 °C 30 μg of λKM4 vector DNA with SpeI (250 U) and NotI (250 U) restriction endonucleases in a final volume of 300 μL (*see* **Note 2**).
5. Purify bacteriophage DNA by using standard phenol/chloroform extraction methods [20].
6. Transfer the purified DNA solution (300 μL) in a 1.5 mL conical microfuge tube, and then add 90 μL of 3 M CH_3COONa pH 5.2, 560 μL of isopropanol, and 510 μL of H_2O. Precipitate the digested DNA by incubating the tube on ice for 20 min.
7. Centrifuge the tube at 16,800 × *g* for 15 min at 4 °C in a benchtop microfuge. Remove the solution, wash the DNA pellet with ice-cold 70% ethanol, and suspend the DNA in 50 μL of 10 mM Tris–HCl pH 8.5 (*see* **Note 3**).
8. Store the purified DNA at −20 °C in small aliquots until use.

3.2 Preparation of Synthetic Adapters

1. Phosphorylate oligonucleotides K185, K187, K189, K192, K194, and K196 by mixing 1 μmol of each oligonucleotide with 8 U of T4 Polynucleotide Kinase, 2 μL of corresponding 10× T4 Polynucleotide Kinase buffer, and 2 μL of ATP in a final volume of 20 μL. Incubate the mixture at 37 °C for 60 min and then at 65 °C for 30 min to stop the enzymatic reaction (*see* **Note 4**).
2. Anneal the phosphorylated/unphosphorylated oligonucleotide pairs by mixing an equal amount of each partner (i.e., 1 μmol) to form six different adapters as the following:

K185 + K186	K187 + K188	K189 + K190
K191 + K192	K193 + K194	K195 + K196

Place the tubes containing the different mixtures in a water bath at 85 °C for 2 min, and then leave them on the bench until room temperature is reached (i.e., 10–20 min).

3. Store the annealed adapters at −20 °C until use.

3.3 Preparation of S. pneumoniae DNA Fragments for Cloning into Lambda Vectors

1. Pick a single colony of *S. pneumoniae*, transfer the colony in 30 mL of TSB, and incubate the culture in a 5% CO_2-enriched atmosphere at 37 °C for 12–16 h.
2. Purify the genomic DNA by using standard methods [23]. Alternatively, the bacterial DNA can be bought from commercial providers such as the American Type Culture Collection (ATCC).
3. Determine the DNA concentration by reading its absorbance at 260 nm with a spectrophotometer.
4. Digest 5–10 μg of the genomic DNA with 1 U of DNase I-(RNase-free). Incubate the genomic DNA and the DNase enzyme for 20 min at 15 °C in a final volume of 100 μL of DNase buffer. Add EDTA at a final concentration of 50 mM to stop the reaction, and put the tube on ice until next step (*see* **Note 5**).
5. Purify digested genomic DNA by using Wizard SV Gel and PCR Clean-Up System (Promega). Elute the DNA in 50 μL of H_2O, and determine the DNA concentration by reading the absorbance at 260 nm with a spectrophotometer.
6. Recess 3′-protruding termini of the digested genomic DNA fragments by using the T4 DNA polymerase. To this aim, mix 2.5 μg of DNA fragments with 9 U of T4 polymerase, 0.5 μg of bovine serum albumin (BSA), and 0.4 mM dNTPs in a final volume of 50 μL of T4 polymerase buffer. Incubate the mixture for 20 min at 12 °C and then for 20 min at 75 °C to stop the reaction.
7. Purify the blunt-end DNA fragments by phenol–chloroform extraction and ethanol precipitation using standard methods [23]. Finally, suspend the DNA pellet in H_2O at a final concentration of 100–200 ng/μL, and store the DNA at −20 °C until next step.
8. Incubate the purified blunt-end genomic DNA fragments with 20-fold molar excess of the six synthetic adapters bringing SpeI or NotI restriction sites (*see* Subheading 3.2) in the presence of the T4 DNA ligase. To this purpose, mix 1 μg of DNA fragments, 30 pmol of each different adapter, 5 μL of 10× T4 ligase buffer, 50 μg of BSA, and 3000 U of the T4 ligase enzyme in a final volume of 50 μL. Perform the ligation reaction overnight at 16 °C, and then inactivate the T4 ligase enzyme for 20 min at 65 °C.

9. Purify the DNA fragments by using the Wizard SV Gel and PCR Clean-Up System, as per the manufacturer's instructions.
10. Proceed with a further purification step of the DNA fragments ligated to the library adapters by using a 2% agarose gel electrophoresis. Under UV lighting, cut out an agarose slice containing the DNA fragments with molecular lengths comprised between 300 and 1000 bp (*see* **Note 6**). Finally, purify the size-selected DNA-adapter fragments from the agarose slice by using the Wizard SV Gel and PCR Clean-Up System, as per the manufacturer's instructions.

3.4 Lambda Display Library Construction

1. Clone the size-selected genomic DNA-adapter fragments into SpeI–NotI digested λKM4 vector (*see* Subheading 3.1). To this aim, mix 0.5 μg of vector DNA, 10–20 ng of DNA fragments, and 200 U of T4 ligase enzyme in a final volume of 10 μL. Perform the ligase reaction overnight at 16 °C, and then inactivate the T4 enzyme for 20 min at 65 °C.
2. Incubate the whole ligase mixture with commercial lambda packaging extracts. Following manufacturer's instructions, calculate the number of extracts which are necessary for packaging the whole ligase reaction (10 μL), then remove the packaging extracts stored in the freezer (−80 °C), and place them on dry ice. Quickly thaw the packaging extracts until the content of the tubes just begins to thaw, and then add the DNA immediately. Stir the tube with a pipet tip to mix well and incubate the tube for 2 h at 22 °C.
3. Add 500 μL of SM buffer to each tube and mix the content gently. Spin briefly to sediment the debris, and store the supernatant containing the phages at 4 °C (supernatant may be stored at 4 °C for up to 1 month).
4. Pick one single colony of *E. coli* BB4 cells (alternatively the LE392 strain can be used) from the master plate, and inoculate the colony in 100 mL of LB medium supplemented with 0.2% maltose and 10 mM $MgSO_4$. Incubate the culture until the bacterial optical density (OD) at 600 nm reaches 1; then centrifuge and suspend the bacterial pellet in 50 mL of SM buffer.
5. Incubate the bacterial cells (50 mL) with the whole phage supernatant (500 μL) containing the packaged phage particles (this procedure will allow an infection rate ≤1 phage per bacterial cell) for 20 min at room temperature.
6. Test the efficiency of the ligation/packaging/infection before plating the whole phage library, and check also the percentage and the size of the DNA inserts cloned into lambda genome (*see* **Note 7** and Fig. 3).

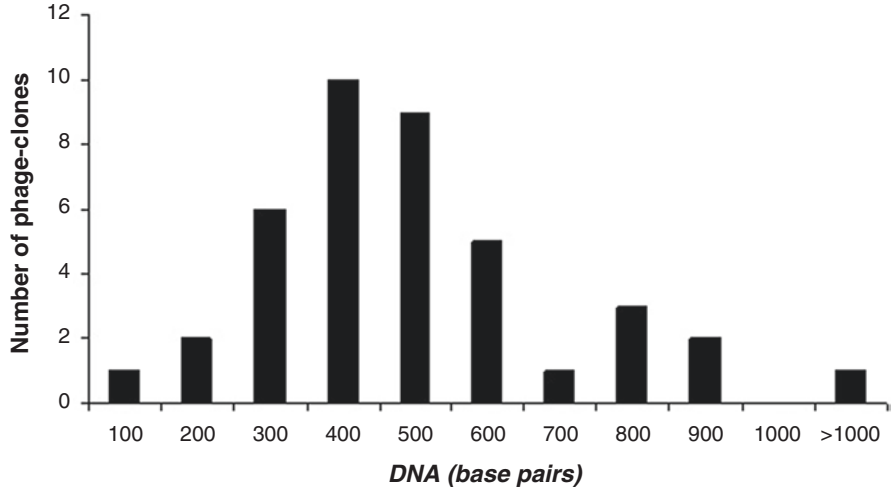


Fig. 3 Distribution of DNA fragments displayed on the surface of phage clones derived from a *S. pneumoniae* gDNA library, as obtained amplifying DNA inserts from single-phage plaques by PCR and analyzing the corresponding products by agarose gel electrophoresis

7. Add 15 mL of top agar medium (pre-warmed at 42 °C) per 0.6 mL of infected cells, and quickly adsorb the mixture to 15 cm NZY-/agar-coated plates (a total of 80–90 plates are needed for one library).
8. Incubate the plates at 37 °C until bacterial plaques are clearly visible (typically 12–16 h).
9. Recover recombinant phages from bacterial plaques by adding 15 mL of SM buffer to each 15 cm plate, and incubate the plates for 4 h at room temperature in an orbital shaker. Transfer the SM buffer containing the eluted phages in conical tubes (50 mL), and remove cell debris by centrifuging the tubes at 3300 × *g* for 20 min at 4 °C.
10. Transfer the supernatant into new tubes, then add NaCl to a final concentration of 1 M, and incubate the tubes for 1 h on ice. After centrifugation at 3300 × *g* for 20 min at 4 °C, collect the supernatant, and add 10% (v/v) of PEG-8000 to precipitate the phage particles.
11. After an overnight incubation at 4 °C, collect the bacteriophage particles by centrifugation at 3300 × *g* at 4 °C for 30 min, and finally suspend the pellet in 1/10 of the starting volume of SM buffer.
12. Repeat the precipitation step with 1 M NaCl and 10% PEG-8000 incubating the tubes on ice for 1 h. Centrifuge the tubes at 3300 × *g* for 30 min at 4 °C, and then suspend the phage pellet in 1/20 of the starting volume of SM buffer.
13. Add DMSO at a final concentration of 7% (v/v). Store the phage suspension in aliquots of 0.5–1 mL at −80 °C.

3.5 Biopanning of Lambda Display Libraries with Human Sera

1. Pick one single colony of *E. coli* BB4 cells from the master plate, and inoculate the colony in 20 mL of LB medium supplemented with 0.2% maltose (w/v) and 10 mM MgSO4. Grow bacterial cells at 37 °C until the bacterial density reaches an $OD_{600\,nm} = 1$; then centrifuge and suspend the bacterial pellet in half volume of SM buffer.
2. Take 100 μL of magnetic beads linked to Protein G (Dynabeads Protein G) for each affinity selection of the lambda library. Wash the beads twice with 1 mL of biopanning coating buffer, and then store the beads on ice in 0.1 mL of coating buffer. Add 10 μL of human serum (i.e., from *S. pneumoniae*-infected individuals) to the washed beads, and incubate the bead suspension with rotation for 40 min at room temperature.
3. Wash the beads three times with 1 mL of biopanning washing buffer, and then incubate the beads with 1 mL of biopanning blocking buffer for 1 h at room temperature.
4. Add 5×10^{10} plaque forming units (pfu) of recombinant lambda phages to the washed beads/serum suspension (final volume of 1 mL of blocking solution). Incubate the mixture for 3–4 h at room temperature with gentle rotation.
5. Wash the beads ten times with 1 mL of biopanning washing buffer, and then add 1.2 mL of bacterial cells (BB4 freshly prepared as described above) to the washed beads/serum/phage mixture. Incubate the mixture for 20 min at room temperature.
6. Add 0.6 mL of the BB4/beads/serum/phage suspension to 15 mL of pre-warmed (42 °C) top agar medium, mix well, and immediately adsorb the mixture onto 15 cm NZY medium-/agar-coated plates.
7. Incubate the plates overnight at 37 °C or until bacterial plaques are clearly visible (typically 12–16 h).
8. Recover recombinant phages from bacterial plaques by adding 15 mL of SM buffer to each 15 cm plate, and incubate the plates for 4 h at room temperature in an orbital shaker. Transfer the SM buffer containing the eluted phages into conical tubes (50 mL), and remove cell debris by centrifuging the tubes at $3300 \times g$ for 20 min at 4 °C.
9. Transfer the supernatant into new tubes, then add NaCl to a final concentration of 1 M, and incubate the tubes for 1 h on ice. After centrifugation at $3300 \times g$ for 20 min at 4 °C, collect the supernatant and add 10% (v/v) of PEG-8000 to precipitate the phage particles.
10. After an overnight incubation at 4 °C, collect the bacteriophages by centrifugation at $3300 \times g$ for 30 min at 4 °C, and finally suspend the pellet in 5 mL of SM buffer.

11. Store the purified phage suspension at 4 °C.
12. Repeat steps 1–10 for the next round of biopanning (usually 2–3 rounds are necessary to sufficiently amplify the phage population selected by the human serum).

3.6 Phage ELISA

1. Coat 96-well MaxiSorp plates with either homemade or commercially available polyclonal anti-lambda antibodies, diluted to 0.1–1 μg/mL in ELISA coating buffer. Incubate overnight at 4 °C, then aspirate the solution, and add 100 μL/well of ELISA blocking buffer. Incubate the plate for 1 h at 37 °C, then remove the blocking buffer, and store the plates without washing at −20 °C until use.
2. Add the recombinant lambda phages selected from the display library to the coated plate (i.e., 10 μL/well of purified phage suspension from biopanning and 90 μL/well of ELISA blocking buffer), and incubate the plate overnight at 4 °C with gentle stirring.
3. Wash the plates using an automated ELISA washer device with ELISA washing buffer. If a manual washing procedure is necessary, wash plate five times with phage ELISA washing buffer and five times with 1× PBS buffer.
4. Dilute 1–2 μL of human serum/plasma (i.e., serum from *S. pneumoniae*-infected individuals) with 10 μL of Fetal Calf Serum (FCS) and 10 μL of BB4 bacterial extract in 100 μL of ELISA blocking buffer (100 μL of final volume). Incubate the mixture for 30 min at room temperature.
5. Add the previous mixtures to the wells and incubate the plate for 1 h at 37 °C with gentle stirring.
6. Wash the plate as in **step 3**.
7. Add 100 μL to each well of antihuman IgG HRP-conjugated antibodies diluted 1:1000 (1 μg/mL) in blocking solution, and incubate the plate for 1 h at 37 °C with gentle stirring.
8. Wash the plates as in **step 3**.
9. Reveal the HRP enzymatic activity by incubating the plates with the chromogenic substrate, TMB, for 10–15 min at room temperature. Stop color development with 25 μL/well of 2 M H_2SO_4.
10. Read the plate by an automated ELISA reader as the difference between the absorbance at 450 and 620 nm (*see* Fig. 4).

3.7 Immunoscreening

1. Inoculate one single colony of *E. coli* BB4 cells in 20 mL of LB medium supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$, until the bacterial density reaches an $OD_{600\ nm} = 2$, and then store the cells on ice.

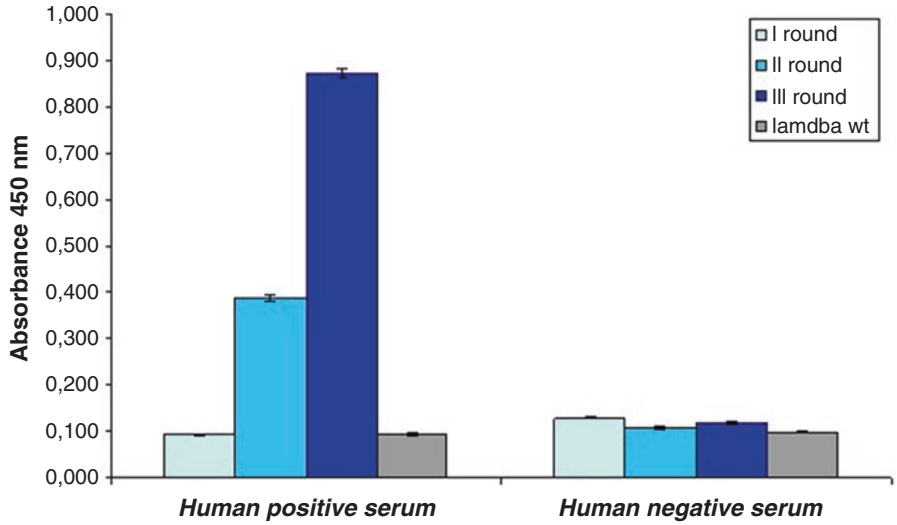


Fig. 4 Phage ELISA reactivity of phage pools after multiple rounds of affinity selection (biopanning). Assayed are pools of phages after biopanning (I, II, and III rounds) and the λKM4 phage (lambda wild type (wt): negative control) with sera from *S. pneumoniae*-infected (positive serum) or uninfected (negative serum) individuals

2. Infect 0.2 mL of bacterial cells with tenfold serial dilutions of recombinant phages (i.e., from 10 pfu/mL up to 10,000 pfu/mL) derived from the affinity selection round (*see* Subheading 3.5), and leave the bacteria/phage mixture for 20 min at room temperature.
3. Add 4–5 mL of pre-warmed top agar medium to the infected cells, and adsorb the resulting mixture immediately onto 90 mm NZY medium-/agar-coated plates. Incubate the plates at 37 °C overnight or until phage plaques are clearly visible.
4. Put 90 mm nitrocellulose filters directly onto the surface of bacterial growth plates. Incubate the plates for 1 h at room temperature, then remove the nitrocellulose filters, and store the plates (representing the master plates) at 4 °C.
5. Incubate the filters with 10 mL of immunoscreening blocking buffer for 1 h at room temperature.
6. Wash twice the filters with 10–15 mL of washing buffer.
7. Incubate the filters with human sera/plasma (i.e., serum from *S. pneumoniae*-infected individuals) diluted 1:100 in immunoscreening blocking buffer supplemented with 10 μL/mL of BB4 extract for 2 h at room temperature.
8. Wash the filters ten times with 10 mL of immunoscreening washing buffer.
9. Incubate the filters with antihuman IgG AP-conjugated antibodies, appropriately diluted (i.e., 1:1000–1:10,000) in immunoscreening blocking buffer for 1 h at room temperature.

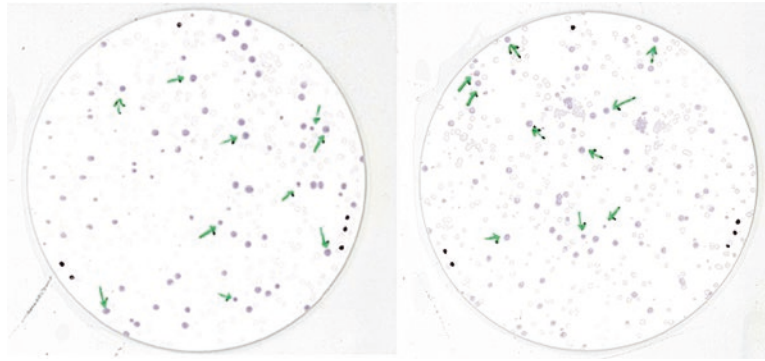


Fig. 5 Immunoscreening of phage pools selected from a *S. pneumoniae* gDNA display library with sera from infected individuals. Nitrocellulose filters are incubated with the same *S. pneumoniae*-positive sera used for library selection, followed by incubation with antihuman IgG alkaline phosphatase-conjugated antibodies. Positive clones are finally revealed with alkaline phosphatase chromogenic substrates. Arrows indicate positive phage clones (dark-gray spots)

10. Reveal the AP enzymatic activity by incubating the filters with the chromogenic substrates (NBT and BCIP substrates) (*see* Fig. 5).

3.8 Antigen Identification Through PCR and DNA Sequencing

1. Overlay the nitrocellulose filters to the master plates, then pick up positive phage plaques (detected by immunoscreening as described in Subheading 3.8) with a sterile tip, and place the phages in a tube by pipetting up and down in 300 μL of SM buffer.
2. Add 50 μL of chloroform and leave the tubes for 15 min at 37 °C. After centrifugation at 16,800 × *g* for 15 min at 4 °C, collect the supernatant and use 2 μL as a template for PCR reaction (final volume of 50 μL) to amplify the recombinant DNA fragments.
3. Use the primers λfor and λrev as specific oligonucleotides mapping, respectively, at N- and C-terminus of the cloning sites in λKM4 vector.
4. Amplify the recombinant DNA fragments by using the following PCR conditions: 30 s at 94 °C, 30 s at 55 °C, and 30 s at 72 °C for 35 cycles.
5. Subject 5 μL of the PCR product to agarose gel electrophoresis (1.5% of agarose gel), and then purify the DNA fragments from the remaining PCR reaction mixture (45 μL) by using Wizard SV Gel and PCR Clean-Up System kits. Store the purified PCR products at −20 °C.
6. Sequence the DNA fragments purified from PCR reactions by using standard methods [20].

7. According to the frame of the DNA fragment with respect to the fusion product with the D gene in λKM4, translate the corresponding nucleotide sequences into protein sequences, and search for sequence similarities between the selected protein fragments and the corresponding sequences in the pathogen gene products (i.e., *S. pneumoniae*) using the available database (i.e., NCBI BLAST Search). This comparison allows the identification of protein fragments containing B-cell epitopes, as being selected by the challenge of the phage library with the antibody repertoire of infected individuals.
8. Use purified PCR products for digesting the DNA fragments with SpeI/NotI restriction endonuclease and cloning the digested fragments into bacterial expression vectors, thus allowing the production, purification, and further characterization of the selected antigen fragments as recombinant proteins, which can be used for diagnosis purposes and vaccine development.

4 Notes

1. The antibodies are not commercially available and should be made using custom services provided by antibody developers (e.g., Pierce, GenScript, etc.) using PEG-purified λKM4 bacteriophages as the antigen.
2. When reaction with endonucleases is simultaneously performed by using two different enzymes, the supplier provides the appropriate reaction conditions (temperature, time, and buffer). In this specific case, SpeI/NotI digestion was carried out in NEB buffer 3 supplemented with 0.1 mg/mL of bovine serum albumin.
3. During preparation and digestion of λKM4 vector, do not use vortex for pellet re-suspension to avoid DNA breaking.
4. Phosphorylation of oligonucleotides representing the 5′-protruding ends of adapters (K185, K187, K189, K192, K194, and K196) is necessary to promote ligation of DNA fragments into λKM4 vector, while the remaining nucleotides (K186, K188, K190, K191, K193, and K195) should not be phosphorylated to avoid ligation between blunt ends of adapters.
5. Before performing the large-scale digestion of the genomic DNA with DNase I, digest a small amount of DNA (100–200 ng), and check the status of DNA fragmentation after different time intervals (i.e., 10, 15, and 20 min of incubation) by agarose gel electrophoresis. If necessary, adjust the length of incubation for the large-scale reaction.
6. When proceeding with the purification of DNA fragments from agarose gels, do not allow the samples to migrate too far

since DNA fragments length should be excised from the gel avoiding to pick up large volume of agarose.

7. Before packaging the whole library, test a small amount of the ligation mixture (i.e., 0.2–0.5 μL), and analyze single-phage clones by PCR and agarose gel electrophoresis (*see* Subheading 3.8) to determine the percentage of cloning and the size of DNA inserts.

References

1. Mikawa YG, Maruyama IN, Brenner S (1996) Surface display of proteins on bacteriophage lambda heads. J Mol Biol 262:21–30
2. Sternberg N, Hoess RH (1995) Display of peptides and proteins on the surface of bacteriophage lambda. Proc Natl Acad Sci U S A 92:1609–1613
3. Niwa M, Maruyama H, Fujimoto T, Dohi K, Maruyama IN (2000) Affinity selection of cDNA libraries by lambda phage surface display. Gene 256:229–236
4. Nicastro J, Sheldon K, Slavcev RA (2014) Bacteriophage lambda display systems: developments and applications. Appl Microbiol Biotechnol 98(7):2853–2866
5. Beghetto E, Pucci A, Minenkova O, Spadoni A, Bruno L, Buffolano W, Soldati D, Felici F, Gargano N (2001) Identification of a human immunodominant B-cell epitope within the GRA1 antigen of Toxoplasma gondii by phage display of cDNA libraries. Int J Parasitol 31:1659–1668
6. Minenkova O, Pucci A, Pavoni E, De Tomassi A, Fortugno P, Gargano N, Cianfriglia M, Barca S, De Placido S, Martignetti A, Felici F, Cortese R, Monaci P (2003) Identification of tumor-associated antigens by screening phage-displayed human cDNA libraries with sera from tumor patients. Int J Cancer 106:534–544
7. Parmley SF, Smith GP (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. Gene 73:305–318
8. Enshell-Seijffers D, Gershoni JM (2002) Phage display selection and analysis of Ab-binding epitopes. Curr Protoc Immunol Chapter 9:Unit 9.8
9. Beghetto E, Spadoni A, Buffolano W, Del Pezzo M, Minenkova O, Pavoni E, Pucci A, Cortese R, Felici F, Gargano N (2003) Molecular dissection of the human B-cell response against Toxoplasma gondii infection by lambda display of cDNA libraries. Int J Parasitol 33:163–173
10. Santini C, Brennan D, Mennuni C, Hoess R, Nicosia A, Cortese R, Luzzago A (1998) Efficient display of an HCV cDNA expression library as C-terminal fusion to the capsid protein D of bacteriophage lambda. J Mol Biol 282:125–135
11. Beghetto E, De Paolis F, Spadoni A, Del Porto P, Buffolano W, Gargano N (2008) Molecular dissection of the human B-cell response against cytomegalovirus infection by lambda-display. J Virol Methods 151:7–14
12. Beghetto E, Gargano N, Ricci S, Oggioni M, Garufi G, Peppoloni S, Pozzi G, Felici F (2006) Discovery of a novel *Streptococcus pneumoniae* antigen by screening a whole genome lambda-display library. FEMS Microbiol Lett 262:14–21
13. De Paolis F, Beghetto E, Spadoni A, Oggioni M, Montagnani F, Felici F, Gargano N (2007) Identification of a human immunodominant B-cell epitope within IgA1 protease of *Streptococcus pneumoniae*. BMC Microbiol 7:113
14. Beghetto E, De Paolis F, Montagnani F, Cellesi C, Gargano N (2009) Discovery of new Mycoplasma pneumoniae antigens by use of a whole-genome lambda-display library. Microbes Infect 11:66–73
15. Montagnani F, De Paolis F, Beghetto E, Gargano N (2010) Use of recombinant chimeric antigens for the serodiagnosis of Mycoplasma pneumoniae infection. Eur J Clin Microbiol Infect Dis 29(11):1377–1386
16. Papasergi S, Lanza Cariccio V, Pietrocola G, Domina M, D'Aliberti D, Trunfio MG, Signorino G, Peppoloni S, Biondo C, Mancuso G, Midiri A, Rindi S, Teti G, Speziale P, Felici F, Beninati C (2013) Immunogenic properties of *Streptococcus agalactiae* FbsA fragments. PLoS One 8(9):e75266
17. Domina M, Lanza Cariccio V, Benfatto S, D'Aliberti D, Venza M, Borgogni E, Castellino F, Biondo C, D'Andrea D, Grassi L, Tramontano A, Teti G, Felici F, Beninati C (2014) Rapid profiling of the antigen regions recognized by serum antibodies using massively parallel sequencing of antigen-specific libraries. PLoS One 9(12):e114159
18. Domina M, Lanza Cariccio V, Benfatto S, Venza M, Venza I, Borgogni E, Castellino F,

Midiri A, Galbo R, Romeo L, Biondo C, Masignani V, Teti G, Felici F, Beninati F (2016) Functional characterization of a monoclonal antibody epitope using a lambda phage display-deep sequencing platform. Sci Rep 6:31458

19. Cano PG, Gamage LNA, Marciniuk K, Hayes C, Napper S, Hayes S, Griebel PJ (2017) Lambda display phage as a mucosal vaccine delivery vehicle for peptide antigens. Vaccine 35(52):7256–7263
20. Fedson DS, Anthony J, Scott G (1999) The burden of pneumococcal disease among adults in developed and developing countries: what is known and what is not known. Vaccine 17:S11–S18
21. Tuomanen EI (2000) Pathogenesis of pneumococcal inflammation: otitis media. Vaccine 19:S38–S40
22. Bogaert D, De Groot R, Hermans PWM (2004) *Streptococcus pneumoniae* colonization: the key to pneumococcal disease. Lancet Infect Dis 4:144–154
23. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY



Chapter 12

Multiplexed Detection of Autoantibodies to Glycopeptides Using Microarray

Johannes W. Pedersen, Alexander Nøstdal, and Hans H. Wandall

Abstract

Protein microarray is a highly sensitive tool for antibody detection in serum. Monitoring of patients' antibody titers to specific antigens is increasingly employed in the diagnosis of several conditions, ranging from infectious diseases, allergies, autoimmune diseases, and cancer. In this protocol, we present a detailed method for enzymatic generation of disease-specific O-glycopeptides and how to monitor the antibody response to these in serum using microarray technology.

Key words Glycopeptide microarray, GalNAc, Sialylation, Serum biomarkers, Autoantibodies, Posttranslational modification, Glycosylation

1 Introduction

Protein microarray enables simultaneous measurement of several biomarkers, e.g., antibodies against foreign antigens or autoantibodies [1, 2]. Antibody profiling has been shown to be useful in differential diagnosis of infectious diseases [3, 4] and allergies [5, 6]. Furthermore, elevated levels of specific autoantibodies are able to identify autoimmune disorders prior to onset of clinical symptoms [7–9]. In the last decades, the autoantibody response to tumor-associated antigens has received much interest for its potential capabilities as a biomarker for early detection of cancer [10]. Compared to the measurement of peptide and protein levels in bodily fluids, antibodies are produced in high, relatively stable titers and have longer half-lives [11]. Besides their use as biomarkers, the identification of antibody targets could reveal specific vaccine candidates, for example, viral and cancer vaccines [12–16].

The discovery of disease-specific antibodies has been approached through several different proteome-wide screening techniques, including expressed cDNA libraries (SEREX) [17], protein and peptide arrays [18, 19], both random and designed phage displays [20], and self-assembling protein arrays [21, 22].

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_12, © Springer Science+Business Media, LLC, part of Springer Nature 2019

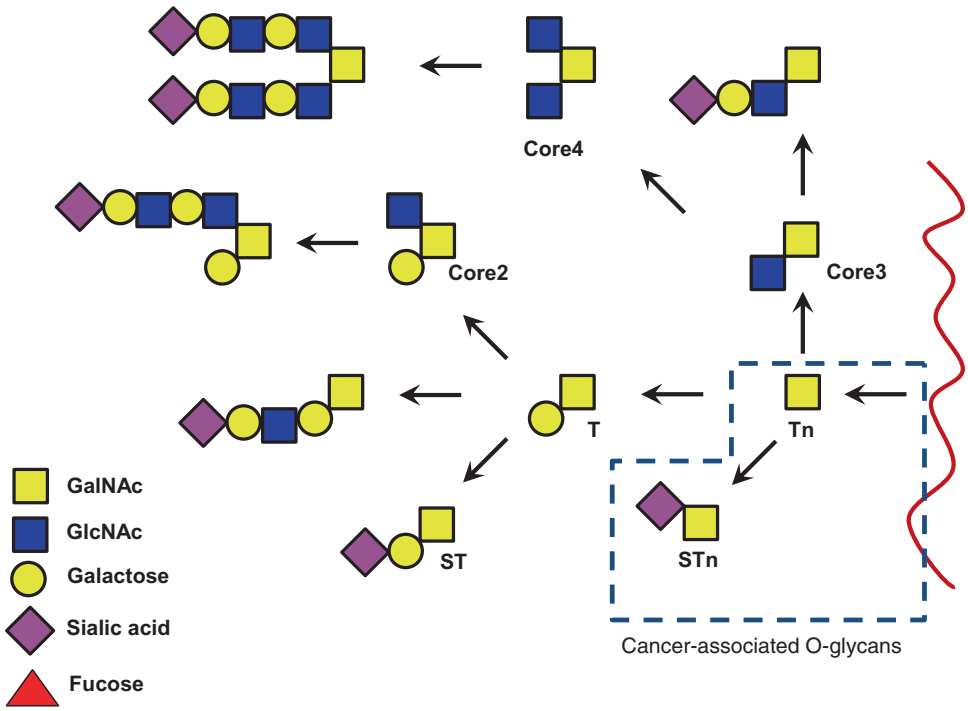


Fig. 1 The O-linked GalNAc glycosylation pathway. Glycosylation is initiated by the addition of GalNAc to serine or threonine residues in the protein core and is elongated through subsequent enzyme reactions. Symbols for monosaccharides GalNAc, Gal, GlcNAc, and Neu5Ac (sialic acid) are indicated

These methods are, however, designed to examine the proteome in the absence of posttranslational modification, which might limit the potential output of antibody screening [14].

Aberrant patterns of O-linked glycosylation (Fig. 1) represent one of the most important cancer-associated posttranslational changes [12, 23]. Changes in O-linked glycosylation are also associated with other pathological conditions, such as autoimmune diseases [24–27] and viral infection [28, 29]. Upregulation of unique short aberrant O-glyco-signatures on proteins may introduce novel glycopeptide epitopes that can elicit autoantibodies because of lack of tolerance.

Such O-glyco-signatures can be displayed on a microarray through printing of pre-synthesized glycopeptides and glycoproteins. The use of pre-synthesized and well-characterized glycopeptides has the advantage of strict control of glycosylation sites and occupancy of the individual glycosylation sites. An alternative high-throughput strategy is in vitro glycosylation directly on peptide or protein arrays. This can be a useful initial screening strategy with multiple antigens. The on-array glycosylation approach can also be combined with on-array in vitro translated proteins in a cell-free system and used in a proteome-wide setting [30].

Pre-synthesized glycopeptides can be produced chemo-enzymatically through in vitro glycosylation of synthetic peptides with the use of recombinant glycosyltransferases. As an alternative, solid-phase peptide synthesis (SPPS) with Fmoc-glycosylated amino acid building blocks can be used for the introduction of initiating carbohydrate moieties at precise location and if needed further elongated in vitro with glycosyltransferases [31, 32]. In this protocol, we will focus on the production of glycopeptide arrays with O-GalNAc glycopeptides synthesized before printing using a chemo-enzymatic approach.

By employing purified glycosyltransferases, we have previously described a method for in vitro O-glycosylation of synthetic MUC1 peptides [13, 14]. The MUC1 protein is a heavily O-glycosylated mucin, predominately in a large 20-amino acid tandem repeat region [33]. While immunological tolerance is experienced to the tandem repeat protein core and its normal glycosylated forms [33, 34], aberrant truncation of the sugar chains results in the induction of autoantibodies to immunodominant O-glycopeptide epitopes. Importantly, the elicited immune response is specifically directed to the combined glycopeptide epitope, with little or no antibody specificity for the Tn carbohydrate hapten. Through strict control of reactive enzymes, we are able to recreate known disease-associated glycosylation patterns on our chosen antigens. In combination with the microarray hydrogel slides, which provide remarkably low background levels [14], this offers a high-throughput method for screening patient sera for antibody reactivity.

In this protocol, we are presenting a chemo-enzymatic approach using a synthetic 60mer tandem repeat MUC1 peptide as example to produce cancer-associated O-glycopeptides. The concept and methods described here could easily be transferred to other kinds of posttranslational modifications. The purpose of this chapter is to provide a detailed protocol for microarray detection of antibodies to glycosylated peptides and to offer aid with troubleshooting based on our experience and studies.

2 Materials

2.1 Generation of O-Glycopeptides

Unless otherwise stated, use deionized water, such as that obtained from a MilliQ water purification system.

2.1.1 Enzymatic GalNAc Glycosylation

1. 5× glycosylation buffer: 125 mM cacodylate and 50 mM manganese(II) chloride ($MnCl_2$) (pH 7.4).

   Make 30 mL 1 M stock cacodylate buffer by dissolving 6.42 g sodium cacodylate trihydrate ($C_2H_6AsO_2Na{\cdot}3H_2O$) in ~20 mL H_2O. Adjust the solution to pH 7.4 using 6 M hydrochloric acid (HCl), and once this is achieved, add H_2O until

total volume reaches 30 mL. Make 30 mL 0.5 M $MnCl_2$ by dissolving 2.96 g $MnCl_2$ in 30 mL H_2O.

Add 3.75 mL 1 M cacodylate pH 7.4 and 3.0 mL 0.5 M $MnCl_2$ to 23.25 mL H_2O, and store at 4 °C, up to 2 months.

2. 100 mM uridine diphosphate-*N*-acetylgalactosamine (UDP-GalNAc).
3. Relevant GalNAc-transferases (GalNAc-Ts): in this protocol we use GalNAc-T1, GalNAc-T2 [35], and GalNAc-T4 [36] to glycosylate our MUC1 peptide. Several trials using different enzymes, with subsequent mass spectrometry (MS) analysis, might be necessary to find a GalNAc-T with a suitable low K_m for the specific glycosylation site(s) (*see* Subheading 3.1 and **Note 1**). Subsequent elongation of GalNAc (Tn)-glycans, as shown in Fig. 1, into NeuAcα2,6GalNAc-S/T (STn) and GlcNAcβ3GalNAc-S/T (Core3) structures requires ST6GalNAc-I [37] and β3GnT6 [38], respectively.
4. Heating incubator (set to 37 °C).

2.1.2 Matrix-Assisted Laser Desorption/Ionization (MALDI) Time-of-Flight (TOF) Mass Spectrometry

1. Buffer A: 0.1% trifluoroacetic acid (TFA), CF_3CO_2H.

   Add 1 mL TFA to 1 L H_2O.
2. Matrix: 10 mg/mL 2,5-dihydroxybenzoic acid (DHB) dissolved in methanol/H_2O (1:1).
3. Matrix-assisted laser desorption/ionization (MALDI) sample plate.
4. Matrix-assisted laser desorption/ionization (Maldi) time-of-flight (TOF) instrument.

2.1.3 Peptide Purification by High-Pressure Liquid Chromatography (HPLC)

1. Buffer A: 0.1% TFA.
2. Buffer B: 90% acetonitrile (ACN), 0.08% TFA.
3. HPLC: Dionex UltiMate 3000.
4. Phenomenex Kinetex 2.6 μM 100 × 4.6 mm c18 column.
5. SpeedVac: Thermo Scientific Savant SPD1010.
6. Labconco Lyph-lock 1 L Lyophilizer.

2.2 Printing

1. Print buffer: 150 mM sodium phosphate (pH 8.5) with 0.005% dimethyl[3-(propyl]azaniumyl]propane-1-sulfonate (CHAPS) and 0.03% sodium azide (NaN_3). Weigh 1.17 g $NaH_2PO_4 \cdot H_2O$, 11.8 g $Na_2HPO_4 \cdot 2H_2O$, and 25 mg CHAPS, and dissolve in H_2O. Add 5 mL 3% NaN_3, and titer the solution to pH 8.5 using 6 M NaOH, and add H_2O until total volume reaches 500 mL.
2. Nexterion® Slide H MPX-48.
3. Light microscope.
4. BioRobotics MicroGrid II spotter.

5. Stealth 3B Micro Spotting Pins.
6. Humidity chamber (*see* **Note 2**).

2.3 Scanning

1. Blocking buffer: 50 mM ethanolamine in 50 mM sodium borate (pH 8.5).

   Weigh 3.05 g ethenolamine and 10.06 g sodium borate decahydrate ($Na_2B_4O_7$), and dissolve in H_2O. Adjust the solution to pH 8.5 using 6 M HCl, and add H_2O until total volume reaches 1 L. Store in amber glass bottle, at room temperature.
2. Phosphate-buffered saline (PBS) pH 7.4.
3. PBS 0.05% Tween (PBS-T) pH 7.4.
4. Cy3-conjugated goat antihuman IgG (Fc-specific) antibodies.
5. Microarray slide centrifuge.
6. ProScanArray HT Microarray Scanner.
7. ProScanArray Express 4.0 software, for image analysis.

3 Methods

3.1 Enzymatic In Vitro O-Glycosylation

In this protocol, we have glycosylated a 60mer MUC1 peptide $(VTSAPDTRPAPGSTAPPAHG)_{n=3}$ representing three tandem repeats, using purified GalNAc-transferases. It is advised that you test your reaction on a small quantity of substrate with several different kinds of GalNAc-transferases, to determine the most suitable transferase(s) (*see* **Note 1**).

More complex glycan structures (e.g., Core3 or STn) can be generated through elongating pre-existing GalNAc residues (Tn antigens), as illustrated in Fig. 2a. The procedures generally follow the same steps as described here, after substituting enzymes and donor sugars. In some cases adjustments to the reaction buffers are necessary (*see* **Notes 3** and **4**).

The quantities used in the protocol are optimized for the glycosylation of 10 μg of the aforementioned peptide (60mer MUC1). Adjustments to the concentrations of UDP-GalNAc, substrate, and enzyme might be necessary for other targets. Likewise the suggested timeframe might need adjustments based on the enzyme's potency and affinity for other targets.

1. Mix 5 μL 5× glycosylation buffer (Subheading 2.1.1) and 1 μL 100 mM UDP-GalNAc in an Eppendorf tube.
2. Add 1 μL substrate (10 μg/μL).
3. Add 0.1 mU GalNAc-T1 and 0.1 mU GalNAc-T2 and GalNAc-T4 (*see* **Note 5**).
4. Add H_2O to a total volume of 25 μL.

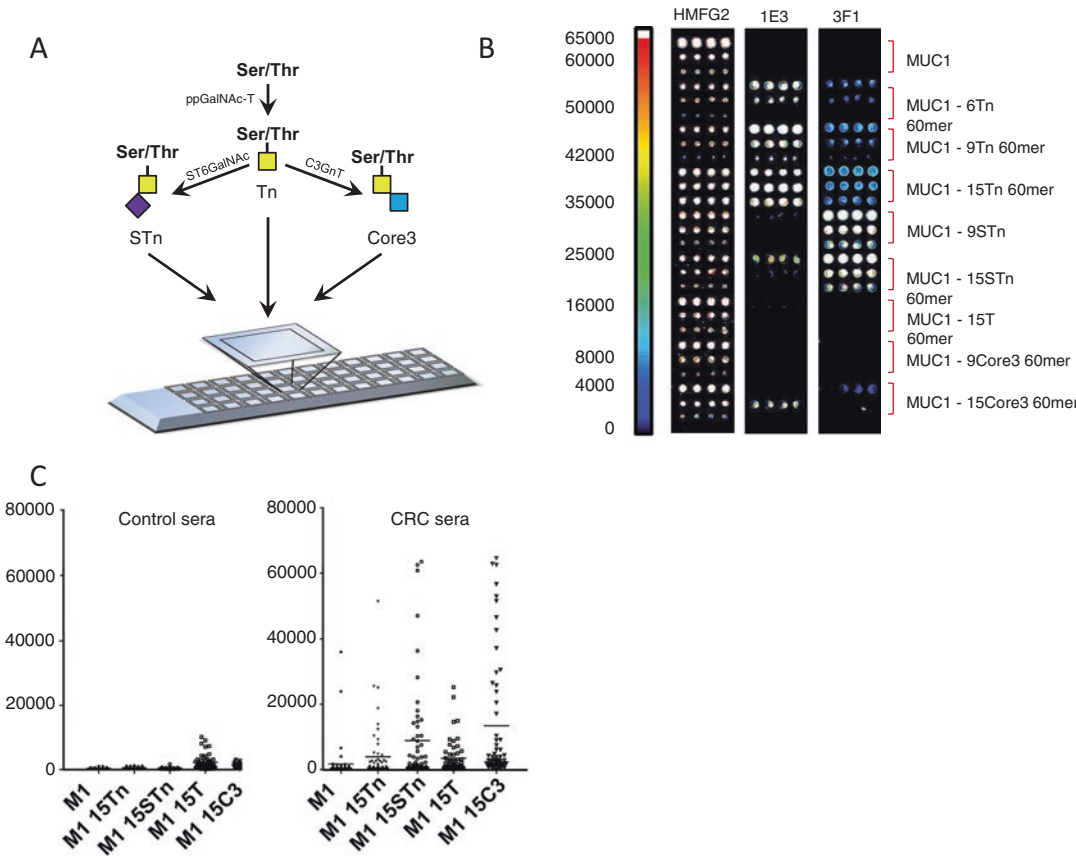


Fig. 2 MUC1 glycopeptides displayed on microarray. (**a**) GalNAc is added to threonine and serine residues on the peptide chain, directed by the specific activity of ppGalNAc-Ts, creating the Tn antigen. The Tn structure can be elongated by subsequent enzyme reactions, e.g., into STn or Core3 structures. The sialyl-Tn antigen is synthesized by the addition of a sialyl residue to GalNAc, whereas addition of a GlcNAc residue to GalNAc results in the formation of the Core3 structure. After purification the glycopeptides are displayed onto a microarray. (**b**) Each peptide is printed in quadruplicates, in three dilutions, and probed with the following antibodies (left to right): MUC1-specific mAb HMFG2, Tn-specific mAb 1E3, and STn-specific mAb 3F1. (**c**) Dot plot diagram presenting IgG autoantibody reactivity against recombinant MUC1, MUC1 Tn, MUC1 STn, and MUC1 Core3. Each dot represents one individual of colorectal cancer patients (serum at time of diagnosis, pretreatment) or healthy controls

5. Vortex mix briefly, and place the Eppendorf tube in the heating incubator at 37 °C, for 12–16 h.
6. Use MALDI-TOF to monitor glycosylation of peptide (Subheading 3.2).
7. STn-MUC1 and Core3-MUC1 are enzymatically synthesized based on purified GalNAc-MUC1 and are characterized and purified as described for GalNAc-MUC1 below (*see* **Notes 3** and **4**).

Note that no detergent (e.g., Triton) is added to the reaction (*see* **Note 6**).

3.2 Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry

1. Dilute 0.5 μL of your sample in 50 μL buffer A (Subheading 2.1.2, **item 1**).
2. Place a 0.5 μL droplet of DHB on the MALDI sample plate. Add 0.5 μL of the diluted enzyme reaction to the DHB.
3. Repeat **steps 1** and **2** of Subheading 3.2 for unglycosylated version of your peptide, for reference.
4. Let the matrix-sample mix dry on the plate; then place the plate in the MALDI-TOF instrument.

Ionize the sample. Use high laser power until a peak correlating to the expected molecular weight appears and then lower laser power to minimize background noise.

After verification of the correct number of GalNAc additions by MALDI-TOF, it is advisable, if possible, to confirm the position of the GalNAc in the intended site(s) [39, 40].

3.3 Peptide Purification by HPLC

Following verification of glycosylation, repeat Subheading 3.1 in larger scale (100–200 μg peptide). Dilute your peptide in buffer A (1:4) (Subheading 2.1.3, **item 1**). Purify the peptide by high-performance liquid chromatography on a C18 column (Subheading 2.1.3, **item 4**), eluting with a 30-min gradient from 10% to 90% ACN in 0.08% TFA (buffer B, Subheading 2.1.3, **item 2**) at 0.700 mL/min flow. Minimize ACN from the samples by using a SpeedVac (Subheading 2.1.3, **item 5**). Lyophilize the sample, and redissolve in H_2O the following day. Verify successful elution of your peptide through MALDI (Subheading 3.2).

3.4 Printing

We employed a BioRobotics MicroGrid II spotter (Genomics Solution) for printing our microarrays. When designing your own microarray, be sure that you make yourself familiar with the hardware and software by reading through the relevant chapters in the user manual. It offers detailed directions by the manufacturer on how to adjust settings to fit your needs.

We recommend printing each compound in triplicates or quadruplets if space allows it on the array and use the mean value of these in the analysis.

3.4.1 Preparing the Sample Source Plate (Biobank)

1. Dilute your non-glycosylated and glycosylated peptides in print buffer (Subheading 2.2, **item 1**). For the first run with new peptides, we recommend running a test print with serial dilutions, to determine the ideal concentration for each compound. Other methods of optimization might be necessary; e.g., highly hydrophobic peptides might benefit from addition of dimethyl sulfoxide (DMSO) in the print buffer, to be fully dissolved (*see* **Note 7**).
2. Seal the biobank with Parafilm to prevent evaporation, and store at −20 °C between runs (*see* **Notes 8** and **9**).

3. Before printing it is important to remove any air bubbles from the wells of the biobank, as they may interfere with the pins' correct acquisition of compounds when dipping into the source plate. Air bubbles are easily removed by centrifugation (e.g., 200 × *g* for 1–2 min).

3.4.2 Printing Peptides

1. We store microarray slides at −20 °C prior to use. Let them equilibrate in room temperature for approximately 30 min before printing; keep them in their sealed envelope for as long as possible. This will prevent condensation on the slides and deactivation.
2. Fill 6 L water reservoir. It might be necessary to refill during the print run, depending on the number of slides, spots per slide, wash settings, and total run time.
3. Examine spotting pins in light microscope. Make sure the pins you choose for your print are clean and unbent. If needed, follow the manufacturer's recommendations for cleaning the pins.

   Place pins in tool, and load it onto the printer.
4. Prime the main wash station. Make sure the water flow is continuous and without air bubbles. To remove air, drain water from the 6 L water reservoir until no air bubbles are visible in the waste tube. Prime the main wash station again, and perform the program "regular wash cycle (without move to bath)" three consecutive times. Repeat if air bubbles persist.
5. Perform one wash cycle (with move to bath); make sure that the pins are lowered into wash station accurately.
6. Load source plate (biobank).
7. Place slides on the tray. Fill out empty slots with spare glass slides. After switching on the vacuum, make sure all slides are aligned correctly.
8. Start the run. The steps for selecting and adjusting run preferences will not be covered here. We advise that PreSpotting and Soft Touch (*see* user manual) are used, for the sake of consistent spot morphology. Make sure the pins are lowered correctly into the source plate wells during their route through the first source visit to spotting and that both PreSpotting and Soft Touch are working as intended.
9. After printing, place the slides in a humidity chamber for 2 h, before you either freeze slides (−20 °C) for later use or proceed directly with serum placement and scanning.

3.5 Serum Placement and Scanning

1. Block slides in blocking buffer (Subheading 2.3, **item 1**) for 1 h. While the slides are being blocked, thaw serum samples on ice. Vortex mix briefly when thawed and before use.
2. Wash slides in PBS 2 × 5 min.

3. Rinse slides in MQ-H_2O and dry through centrifugation (200 × *g*). This ensures that the Teflon coating separating each well will be capable of upholding its function.
4. Place slides in a petri dish.
5. Dilute the sera 1:5 in PBS-T, and add 9 μL to each well of the slide. This can also be done on slide, by placing a 7.5 μL droplet of PBS-T in each well, and then add 1.5 μL undiluted sera (*see* **Note 10**).
6. The addition of Tween to the diluted sera prevents nonspecific binding of autoantibodies.

   When doing large batches, we recommend devoting 1–4 wells per slide to the same sera or antibodies against the printed targets, to easily enable discovery of eventual slide-to-slide variation. This also ensures that in the case of scanner malfunction or service, scanning parameters can be calibrated using the control wells.
7. Let serum incubate on slide for 1 h, under gentle agitation (~50 motions per minute).
8. Wash off the sera: wash with PBS-T for 2 × 5 min and then a third 5-min wash cycle where PBS-T is swapped for PBS after 2.5 min.
9. While the slides are washed, prepare a clean petri dish for the next steps, inside a light-blocking medium (e.g., a box or a sheet of aluminum foil).
10. Dry the edges of the slide off with a tissue, creating a dry Teflon frame around the wells.
11. Dilute Cy3-conjugated goat antihuman IgG (Fc-specific) antibodies 1:4000 in MQ-H_2O.

    Place 1 mL per slide inside the dried Teflon frame and incubate for 1 h.

    The slides should henceforth be shielded from light, by closing the box/aluminum foil around the petri dish.
12. Wash off the goat antihuman antibodies: wash with PBS-T for 2 × 5 min and then a third 5-min wash cycle with PBS. The slides should be shielded from light during these washes.
13. Rinse slides in H_2O, dry by centrifugation (200 × *g*), and place slide in scanner cassette.

    Scan the slides using ProScanArray HT Microarray Scanner (PerkinElmer). We recommend scanning with 100% laser power at several different PMT gains (typically ranging from 70% to 90%) to find a suitable standard for your samples, with reasonably low background levels and where the reactive spots' intensity is easily readable. Resolution of the images should not surpass 10 μm (*see* **Note 11**).

14. After scanning, seal slides in a container by Parafilm, and freeze at −20 °C. If needed, slides can be thawed and rescanned (e.g., at other scanning parameters).

3.6 Quantification and Data Analysis

After scanning, set up a quantification protocol as per the Scan Array Express user manual. To simplify the data analysis, we highly recommend the creation and use of a. GAL file for the quantification template, so that the name and ID tag for each printed compound is imported automatically to their corresponding spots. By following the printing specifications used in this protocol, each spot should average 150 μm in diameter. When the quantification template is correctly aligned on top of the scanned image of the microarray, export the result spreadsheet to excel, and convert the text to columns.

Each spot is now easily identifiable with correct name; and the array column and row information enables sorting of the data after serum sample. The final data output is gained from the mean spot relative fluorescence units (RFU) after subtracting the surrounding background (found in column Y in the Excel spreadsheet, labeled as "Ch1 Mean – B") (*see* **Note 12**).

4 Notes

1. There are over 200 glycosyltransferases in human cells, divided into 90 distinct families sharing mutual donor-sugar substrates. Of the GalNAc-transferase (GalNAc-T) family, there are 20 known members, with varying affinity and specific activity for each potential glycosylation site. Although some effort has been put into mapping glycosylation sites in different proteins for specific GalNAc-Ts, some trial and error employing different enzymes is usually needed to get the desired sites glycosylated. Also, if a peptide holds more than one potential site, it is not uncommon that several different transferases are needed to glycosylate them all.
2. Commercial humidity chambers are available; however, we made our own using a sealable plastic container with a perforated tray for slides. The lower compartment was filled with NaCl-saturated water, stabilizing the relative humidity at ~75%, room temperature.
3. The purified Tn-bearing peptide can be enzymatically sialylated with ST6GalNAc-I (sialyl transferase), in the following reaction mixture: 50 mM MES (pH 6.5), 20 mM EDTA, and 2 mM DTT. The donor sugar used for this reaction is CMP-NeuAc, at 2 mM.
4. Enzymatic elongation of the purified Tn-bearing peptide can be performed by β3Gn-T6 (Core3 synthase), in the following

reaction mixture: 50 mM cacodylate (pH 7.4) and 10 mM $MnCl_2$. The donor sugar used for this reaction is UDP-GlcNAc, at 2 mM.

5. The amounts of enzyme (0.1 mU GalNAc-T1, 0.1 mU GalNAc-T2, and 0.1 mU GalNAc-T4) were determined by scintillation counting after Dowex-1 formic acid chromatography of 25 μL reaction assays as described in Subheading 3.1, with UDP-GalNAc substituted for UDP-[$_{14}$C]GalNAc (2000 cpm/nmol) (Amersham Biosciences). For GalNAc-T1 and GalNAc-T2, IgA-hinge peptide (VPSTPPTPSPSTPPPTSPSK) was used for the activity assay; for GalNAc-T4, a MUC7 peptide (APPTPSATTPAPPSSSAPPETTAA) was used.
6. Avoid using detergent during enzymatic O-glycosylation with soluble glycosyltransferases when possible. It is difficult to completely remove detergent during purification, and this could later cause problems when printing. Compounds mixed with detergent will often create nonhomogeneous spot morphologies and may in worst case cause spots to merge, rendering slide uninformative and unusable.
7. Very high concentrations of DMSO are not recommended for printing, as it may interfere with the hydrogel surface of the microarray slides; we have, however, successfully printed with up to 60% DMSO in our print buffer.
8. During long runs, water evaporation of the biobank will lower the volume of each sample and increase the concentration in each well. In these cases, it is possible to refill the wells with MQ-H_2O, before a new print.
9. Precipitation issues may arise after thawing a frozen biobank. A way to solve this problem is to incubate the thawed biobank at 37 °C for 20 min before printing. A more laborious method is to manually pipette up and down repeatedly in the well of each compound, thus homogenizing the content of each well.
10. Dilution of sera is application and source dependent and should be optimized for each experimental design. In the current protocol, we use a serum dilution of 1:5.
11. As with the microarray printer hardware and software, it is important that you have made yourself familiar with the scanner and its software. The manufacturer's user manual offers a great overview on how to use the scanner, and after reading the relevant chapters in the manual, it is fairly easy to adjust the default settings to those specified in this protocol.
12. Presumably due to varying content composition between different serum samples, it is not unusual to observe varying levels of background signal between subarrays of a microarray slide, directly corresponding to the area covered by serum.

To remove this variance from the results, the background level is automatically measured by the software and subtracted from the mean spot RFU.

References

1. Kingsmore SF (2006) Multiplexed protein measurement: technologies and applications of protein and antibody arrays. Nat Rev Drug Discov 5(4):310–320
2. Robinson WH (2006) Antigen arrays for antibody profiling. Curr Opin Chem Biol 10(1):67–72
3. Chevaliez S, Pawlotsky JM (2008) Diagnosis and management of chronic viral hepatitis: antigens, antibodies and viral genomes. Best Pract Res Clin Gastroenterol 22(6):1031–1048
4. Murdoch DR et al (2012) Laboratory methods for determining pneumonia etiology in children. Clin Infect Dis 54(Suppl 2):S146–S152
5. Fall BI et al (2003) Microarrays for the screening of allergen-specific IgE in human serum. Anal Chem 75(3):556–562
6. Ferrer M et al (2009) Molecular diagnosis in allergology: application of the microarray technique. J Investig Allergol Clin Immunol 19(Suppl 1):19–24
7. Leslie D, Lipsky P, Notkins AL (2001) Autoantibodies as predictors of disease. J Clin Invest 108(10):1417–1422
8. Routsias JG, Tzioufas AG, Moutsopoulos HM (2004) The clinical value of intracellular autoantigens B-cell epitopes in systemic rheumatic diseases. Clin Chim Acta 340(1–2):1–25
9. Wu J et al (2017) The roles and applications of autoantibodies in progression, diagnosis, treatment and prognosis of human malignant tumours. Autoimmun Rev 16(12):1270–1281
10. Tainsky MA (2009) Genomic and proteomic biomarkers for cancer: a multitude of opportunities. Biochim Biophys Acta 1796(2): 176–193
11. Anderson KS, LaBaer J (2005) The sentinel within: exploiting the immune system for cancer biomarkers. J Proteome Res 4(4): 1123–1133
12. Tarp MA, Clausen H (2008) Mucin-type O-glycosylation and its potential use in drug and vaccine development. Biochim Biophys Acta 1780(3):546–563
13. Pedersen JW et al (2010) Seromic profiling of colorectal cancer patients with novel glycopeptide microarray. Int J Cancer 128(8): 1860–1871
14. Wandall HH et al (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res 70(4):1306–1313
15. Clo E et al (2012) Characterization of the viral O-glycopeptidome: a novel tool of relevance for vaccine design and serodiagnosis. J Virol 86(11):6268–6278
16. Pedersen JW et al (2014) Cancer-associated autoantibodies to MUC1 and MUC4—a blinded case-control study of colorectal cancer in UK collaborative trial of ovarian cancer screening. Int J Cancer 134(9):2180–2188
17. Sahin U et al (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. Proc Natl Acad Sci U S A 92(25):11810–11813
18. Pereira-Faca SR et al (2007) Identification of 14-3-3 theta as an antigen that induces a humoral response in lung cancer. Cancer Res 67(24):12000–12006
19. Stockert E et al (1998) A survey of the humoral immune response of cancer patients to a panel of human tumor antigens. J Exp Med 187(8):1349–1354
20. Mintz PJ et al (2003) Fingerprinting the circulating repertoire of antibodies from cancer patients. Nat Biotechnol 21(1):57–63
21. Anderson KS et al (2008) Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in breast cancer. J Proteome Res 7(4):1490–1499
22. Ramachandran N et al (2004) Self-assembling protein microarrays. Science 305(5680):86–90
23. Schietinger A et al (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen. Science 314(5797):304–308
24. Chui D et al (2001) Genetic remodeling of protein glycosylation in vivo induces autoimmune disease. Proc Natl Acad Sci U S A 98(3):1142–1147
25. Opdenakker G et al (2006) Remnant epitopes, autoimmunity and glycosylation. Biochim Biophys Acta 1760(4):610–615
26. Axford JS et al (1992) Changes in normal glycosylation mechanisms in autoimmune rheumatic disease. J Clin Invest 89(3):1021–1031

27. Delves PJ (1998) The role of glycosylation in autoimmune disease. Autoimmunity 27(4): 239–253
28. Olofsson S, Blomberg J (1977) Studies on glycopeptides of Herpes simplex virus infected cells. Arch Virol 55(4):293–304
29. Brennan PJ et al (1976) Metabolism of galactose in herpes simplex virus-infected cells. Virology 69(1):216–228
30. Karthikeyan K et al (2016) A contra capture protein array platform for studying post-translationally modified (PTM) auto-antigenomes. Mol Cell Proteomics 15(7):2324–2337
31. Herzner H et al (2000) Synthesis of glycopeptides containing carbohydrate and peptide recognition motifs. Chem Rev 100(12): 4495–4538
32. Blixt O, Clo E (2013) Synthesis of O-glycopeptides and construction of glycopeptide microarrays. Methods Mol Biol 1047: 201–214
33. Taylor-Papadimitriou J et al (2002) MUC1 and the immunobiology of cancer. J Mammary Gland Biol Neoplasia 7(2):209–221
34. Rughetti A et al (2005) Recombinant tumor-associated MUC1 glycoprotein impairs the differentiation and function of dendritic cells. J Immunol 174(12):7764–7772
35. Wandall HH et al (1997) Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3. J Biol Chem 272(38):23503–23514
36. Bennett EP et al (1998) Cloning of a human UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. J Biol Chem 273(46):30472–30481
37. Ikehara Y et al (1999) Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. Glycobiology 9(11):1213–1224
38. Iwai T et al (2002) Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide beta1,3-N-acetylglucosaminyltransferase (beta 3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans. J Biol Chem 277(15):12802–12809
39. Mirgorodskaya E et al (1999) Partial vapor-phase hydrolysis of peptide bonds: a method for mass spectrometric determination of O-glycosylated sites in glycopeptides. Anal Biochem 269(1):54–65
40. Pedersen JW et al (2011) Lectin domains of polypeptide GalNAc transferases exhibit glycopeptide binding specificity. J Biol Chem 286(37):32684–32696



Chapter 13

Construction and Screening of an Antigen-Derived Peptide Library Displayed on Yeast Cell Surface for CD4+ T Cell Epitope Identification

Fei Wen, Mason R. Smith, and Huimin Zhao

Abstract

Antigenic peptides (termed T cell epitopes) are assembled with major histocompatibility complex (MHC) molecules and presented on the surface of antigen-presenting cells (APCs) for T cell recognition. T cells engage these peptide-MHCs using T cell receptors (TCRs). Because T cell epitopes determine the specificity of a T cell immune response, their prediction and identification are important steps in developing peptide-based vaccines and immunotherapies. In recent years, a number of computational methods have been developed to predict T cell epitopes by evaluating peptide-MHC binding; however, the success of these methods has been limited for MHC class II (MHCII) due to the structural complexity of MHCII antigen presentation. Moreover, while peptide-MHC binding is a prerequisite for a T cell epitope, it alone is not sufficient. Therefore, T cell epitope identification requires further functional verification of the MHC-binding peptide using professional APCs, which are difficult to isolate, expand, and maintain. To address these issues, we have developed a facile, accurate, and high-throughput method for T cell epitope mapping by screening antigen-derived peptide libraries in complex with MHC protein displayed on yeast cell surface. Here, we use hemagglutinin and influenza A virus X31/A/Aichi/68 as examples to describe the key steps in identification of CD4+ T cell epitopes from a single antigenic protein and the entire genome of a pathogen, respectively. Methods for single-chain peptide MHC vector design, yeast surface display, peptide library generation in *Escherichia coli*, and functional screening in *Saccharomyces cerevisiae* are discussed.

Key words CD4+ T cell epitope mapping, Peptide library, Major histocompatibility complex, MHC-binding peptides, Human leukocyte antigen DR1 (HLA-DR1), Yeast display, Influenza A virus, Flow cytometry, High-throughput screening

1 Introduction

T cells respond to selected peptides (termed T cell epitopes) in complex with major histocompatibility complex (MHC) molecules expressed on antigen-presenting cells (APCs) through their unique surface receptors (T cell receptors (TCRs)) [1, 2]. T cell epitopes play a critical role in T cell development and activation, and their identification has long been important for studying T cell lineages

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_13, © Springer Science+Business Media, LLC, part of Springer Nature 2019

and phenotypes [3, 4]; elucidating self-tolerance mechanisms [5], vaccine design, and assessment [6–8]; tracking T cells in vivo [9]; etc. More recently, T cell epitope identification has been shown to be an important aspect of cancer immunotherapy, and several tumor-associated T cell epitopes have been identified and leveraged to produce positive therapeutic outcomes [10–12]. However, as tumors grow, cancer cells expressing common tumor-associated T cell epitopes are cleared via T cell-dependent immunoselection, promoting the expansion of cancer cells lacking known T cell epitopes [13]. As a result, considerable effort has been devoted to predicting and identifying new patient-specific T cell epitopes derived from unique tumor neoantigens [14, 15].

Within the last two decades, a number of diverse computational methods have been developed to predict possible T cell epitopes by evaluating if a peptide will bind a particular MHC. These approaches include matrix-based methods [16, 17], structure-based methods [18, 19], and machine learning [20–22]. We have recently shown that structure-based molecular dynamic simulations using a weighted histogram analysis method can accurately predict peptide-MHCII binding affinity; however, this method is computationally intensive limiting its use in screening large protein libraries [19]. The accuracy of most computational peptide-MHC binding predictions is measured by the area under the curve (AUC) of the receiver operating characteristic (ROC). Machine learning methods are among the most accurate to date, with computational peptide-MHC class I (MHCI) binding predictions achieving an AUC of 0.85–0.95 and peptide-MHCII binding predictions achieving an AUC of 0.75–0.85 [23]. The much lower prediction accuracy for MHCII-binding peptides is mainly due to the structural complexity of MHCII antigen presentation. For example, the open-ended nature of the MHCII peptide-binding groove allows for the presentation of peptides that are longer (13–25 amino acids) and of more variable length than MHCI. Moreover, flanking residues on long peptides bound to MHCII can interact with peripheral residues on the MHCII in ways that are difficult to predict, which may affect peptide-MHCII binding.

Despite these complicating factors, peptide-MHCII binding prediction accuracy can be improved by combining top-performing individual methods into a consensus prediction [24, 25], which is the method recommended by the Immune Epitope Database (IEDB) [26]. While these methods perform well during cross validation with standardized datasets, they tend to underperform when applied to new datasets or datasets containing peptides of different lengths [22]. In a comprehensive study of 21 different computational methods for predicting peptide-MHCII binding, the most reliable methods could only predict 50% of T cell epitopes within four antigenic protein libraries, and all were associated with high false-positive rates [27].

Given the limited accuracy and high false-positive rates associated with computational methods for predicting peptide-MHCII binding, robust experimental methods are necessary for reliable CD4+ T cell epitope identification. The most straightforward T cell epitope identification method is to use HPLC-MS to determine the sequence of the peptides extracted directly from the peptide-MHC (pMHC) complexes on the surface of professional APCs [28]. However, this method has met limited success due to the relatively small number of peptides attainable and representation of antigenic peptides as nested sets with heterogeneous lengths. When the antigenic protein sequence is known, the most common method for T cell epitope identification is to chemically synthesize overlapping peptides spanning the entire protein and load these peptides individually onto professional APCs (usually irradiated peripheral blood mononuclear cells) or artificial APCs. The peptide-loaded APCs are then tested for their ability to activate T cells [29–31]. To this end, we have recently demonstrated that yeast cells engineered to present human MHCII can serve as a modular platform for rapidly assessing T cell activation [32]. Although effective for limited studies, the individual investigation of overlapping peptide libraries becomes impractical when the overall size of the antigenic proteins being screened becomes large (e.g., multiple proteins in viruses with large genome size) and is not applicable when the antigen sequence is unknown. As an alternative and more comprehensive approach, combinatorial synthetic peptide libraries with up to a trillion variants, such as positional scanning synthetic combinatorial libraries (PS-SCLs) [33] and bead-bound libraries [34], can be employed. These synthetic combinatorial peptide libraries are especially valuable when the relevant target antigen is unknown. Nevertheless, these methods rely on chemical synthesis of peptides, which is generally expensive and nonrenewable. As a result, several cloning strategies have been developed to generate deoxyribonucleic acid (DNA) libraries encoding antigenic proteins.

To directly identify T cell epitopes, several groups have developed methods by either expressing a complementary DNA (cDNA) library in engineered APCs [35] or expressing viral cDNA libraries in *E. coli* to identify the antigenic protein first [36], so that the epitopes could be identified by testing chemically synthesized overlapping peptides. These expression cloning methods usually require either extensive engineering of a cell line for efficient processing and presenting MHCII-restricted antigens that are usually of exogenous origin or the use of professional APCs that are often difficult to isolate, expand, and maintain. Therefore, the screening process is often time-consuming, laborious, and reagent-intensive.

To address some of the limitations discussed above, we have developed a facile and high-throughput CD4+ T cell epitope mapping method by displaying pathogen-derived peptide libraries in complex with the restriction MHCII protein on yeast cell surface

[37]. It presents the advantages of both combinatorial peptide libraries and expression cloning and allows direct epitope identification from either known or unknown pathogens. Here, we use hemagglutinin and influenza virus X31/A/Aichi/68 as an example for each case. Both antigens contain $HA_{306–318}$ peptide (PKYVKQNTLKLAT) that has been well characterized in complex with the human MHCII allele, DR1. The design of the epitope mapping method is shown in Fig. 1. DNA encoding the pathogenic protein(s) is obtained by polymerase chain reaction (PCR) or reverse transcription-PCR (RT-PCR) using sequence-specific or random hexamer primers (Fig. 1a) for known and unknown antigens, respectively. For RT-PCR, RNase H and DNA polymerase I are used to generate the second-strand cDNA. The resulting double-stranded DNA is then randomly digested by DNase I into fragments with size ranging from 30 to 60 nucleotides that encode all the possible peptides with 10–20 amino acids from the antigens, which are then purified and blunt-end polished (T4 DNA polymerase and Klenow fragment). Meanwhile, the yeast display vector containing the gene encoding the restriction single-chain MHCII molecule, e.g., DR1 (Fig. 1b), is digested with *Not*I/*Spe*I (*see* **Note 1**), blunt-end polished, and then ligated to the DNA fragments (Fig. 1c). Following gene expression, yeast cells displaying the peptide library in complex with DR1 are analyzed by fluorescence-activated cell sorting (FACS) to enrich those containing peptides with high DR1-binding affinity based on their surface expression levels (*see* **Note 2**). This enrichment cycle is repeated for ~2 to 4 times to obtain a small subset of yeast cells displaying DR1 in complex with high-affinity binding peptides. The selected DR1-bound peptides are then directly screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 secretion using T hybridoma cells transfected with the TCR of interest (*see* **Note 3**). DNA sequence analysis of the selected positive clones leads to the identification of antigenic peptides. If necessary, a series of DNA fragments encoding overlapping peptide sequences can be used to refine those antigenic peptide sequences.

2 Materials

All buffers and media were prepared using ultrapure water (ddH_2O) with a resistivity of 18.2 mΩ cm at room temperature and were sterilized by autoclaving at 121 °C and 15 psi for 15 min or filtration through a 0.22 μm membrane. All restriction enzymes were obtained from New England Biolabs and stored at −20 °C, and all antibodies were stored at 4 °C. All primers (listed in Table 1) were synthesized by Integrated DNA Technologies and stored at −20 °C at a concentration of 100 μM. All other reagents were stored at room temperature unless indicated otherwise.

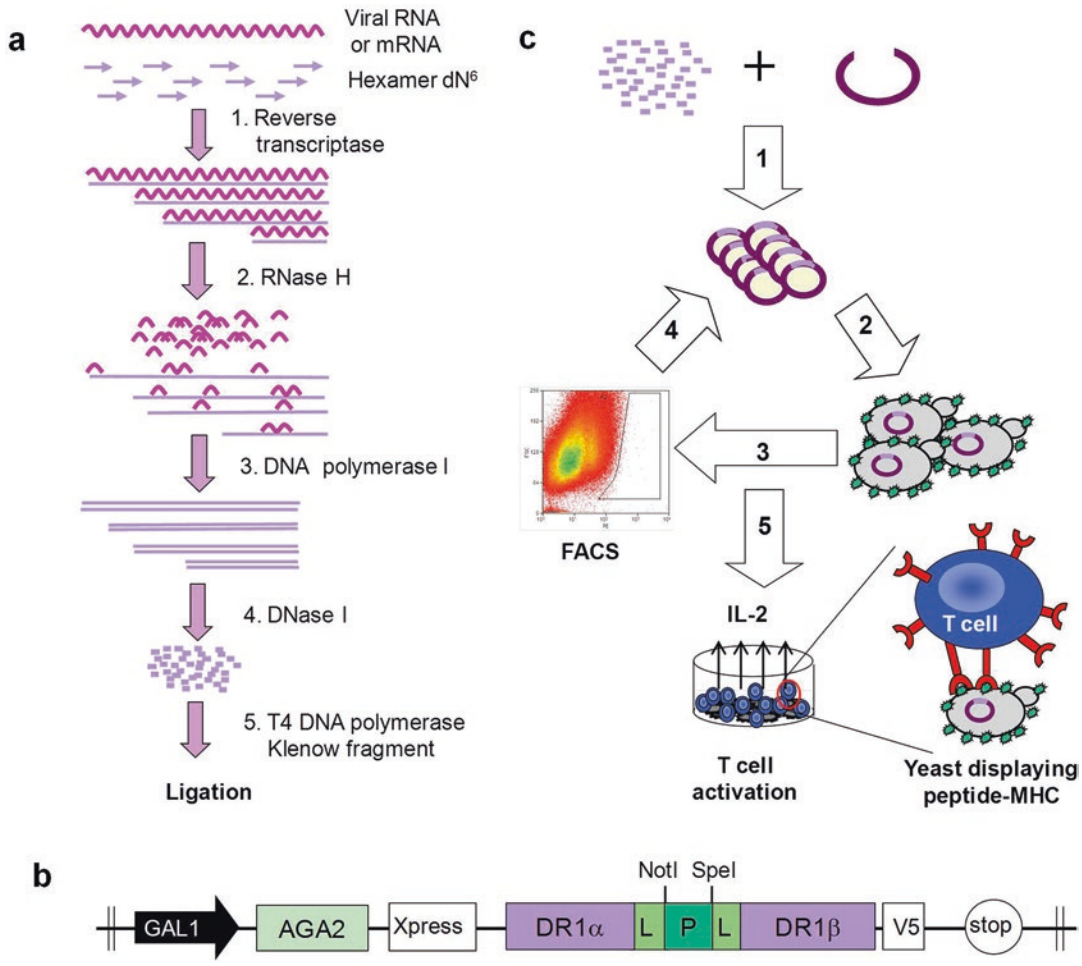


Fig. 1 Design of the CD4+ T cell epitope mapping method. (**a**) Generation of blunt-ended DNA fragments (~30 to 60 bp) encoding pathogen-derived peptides from genomic RNA or mRNA isolated from pathogens with unknown sequence. For antigens with known sequences, double-stranded DNA could be readily amplified using specific primers by PCR and processed from step 4 with DNase I treatment. (**b**) Schematic representation of the single-chain peptide-DR1 construct used for yeast display. ***GAL1*** yeast GAL1 promoter, ***AGA2*** an adhesion subunit of a-agglutinin of a-yeast cells [43], ***Xpress*** Xpress epitope, ***L*** linker, ***P*** peptide, ***V5*** V5 epitope. (**c**) Peptide library generation and screening. DNA fragments are ligated to the single-chain MHCII molecules in a yeast display vector (step 1). Following transformation and gene expression (step 2), yeast cells displaying the peptide library in complex of MHCII protein are analyzed by FACS (step 3) to identify a small subset of yeast cells containing peptides with high affinity toward the restriction MHCII. Plasmids are then recovered from these yeast cells (step 4) and analyzed for further enrichment. This enrichment cycle is usually repeated for ~2 to 4 rounds. Individual clones from the enriched library are then screened for their ability to induce antigen-specific T cell activation as indicated by IL-2 production using T cell hybridomas transfected with the TCR of interest (step 5) in a 96-well format

Table 1
Primers used in cloning and peptide library creation

Name	Sequence
α-5BX	5′-GTACCAGGATCCAGTGTGGTGGAAAGGAAAGAAGAACATGTGATC-3′
β-3XH	5′-CCCTCTAGACTCGAGCTTGCTCTGTGCAGATTCAGAC-3′
pYD1For	5′-AGTAACGTTTGTCAGTAATTGC-3′
NotIRev	5′-TGCCAACTTCAGGGTGTTTTGCTTAACATACTTGGGGCGGCCGCCTCCTGAGCCTCCACC-3′
NotIFor	5′-GCAAAACACCCTGAAGTTGGCAACAGGTACCGGTGGCTCACTAG-3′
βRev73-67	5′-GGCCCGCCTCTGCTCCAGGA-3′
StfFor	5′-GGAGGCGGCCGCTTT TTG GATGGAGGAATTCATATG-3′
StfRev	5′-CTCACTAGTCGGGAAGACGTACGGGGTATACATGT-3′
AichiFor	5′-ATTCGCGGCCGCATGAAGACCATCATTGCTTTGAGCTACATTTTC-3′
AichiRev	5′-CTAATAACTAGTAATGCAAATGTTGCACCTAATGTTGCCTCTCTG-3′
2467For	5′-GGCCGCCCCAAGTATAGAAAGATGAACGCACGAAAGTTGGCAACAGGTACCGGTGGCTCA-3′
2467Rev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTTCGTGCGTTCATCTTTCTATACTTGGGGC-3′
M4For	5′-GGCCGCGGAGGTTATAGACAGATGTCAGCACCAACTTTGGGAGGCGGTACCGGTGGCTCA-3′
M4Rev	5′-CTAGTGAGCCACCGGTACCGCCTCCCAAAGTTGGTGCTGACATCTGTCTATAACCTCCGC-3′
YAKFor	5′-GGCCGCGCCGCATATGCCGCAGCGGCTGCCGCAAAGGCTGCCGCAGGTACCGGTGGCTCA-3′
YAKRev	5′-CTAGTGAGCCACCGGTACCTGCGGCAGCCTTTGCGGCAGCCGCTGCGGCATATGCGGCGC-3′
CIIFor	5′-GGCCGCGCTGGGTTTAAGGGGGAACAGGGACCTAAAGGAGAGCCTGGTACCGGTGGCTCA-3′

CIIRev	5′-CTAGTGAGCCACCGGTACCAGGCTCTCCTTTAGGTCCCTGTTCCCCCTTAAACCCAGCGC-3′
PKAFor	5′-GGCCGCCCCAAGGCTGTTAAGCAAAACACCCTGAAGTTGGCAACAGGTACCGGTGGCTCA-3′
PKARev	5′-CTAGTGAGCCACCGGTACCTGTTGCCAACTTCAGGGTGTTTTGCTTAACAGCCTTGGGGC-3′
LFor	5′-GGCCGCGGAGGTGGAGGCTCCGGAGGTGGAGGCTCAGGAGGTGGAGGTACCGGTGGCTCA-3′
LRev	5′-CTAGTGAGCCACCGGTACCTCCACCTCCTGAGCCTCCACCTCCGGAGCCTCCACCTCCGC-3′
MBPNotI	5′-GGAGGCGGCCGCGAAAACCCGGTTGTTCACTTCTTCAAAAACATCGTTACCCCGCGTGGTACCGGTGGCTCACTAGTGA-3′

2.1 Vector Construction and DNA Fragment Generation

1. pJ3/238: A gift from M. Mage (NIH, Bethesda, MD) was used as the PCR template for amplifying the DNA encoding a single-chain polypeptide consisting of the α-chain of DR1, the peptide $HA_{306–318}$, and the β-chain of DR1 (Fig. 1b).
2. 5× Phusion HF reaction buffer and Phusion DNA polymerase.
3. 40× deoxynucleotide triphosphate (dNTP) premix: 10 mM of each nucleotide.
4. 0.5 M ethylenediaminetetraacetic acid (EDTA) solution (pH 8.0): For a 500 mL of stock solution of 0.5 M EDTA, weigh out 93.05 g of EDTA disodium salt (molecular weight (MW) = 372.2 g/mol), and dissolve it in 400 mL of deionized water. Adjust to pH 8.0 with NaOH and correct the final volume to 500 mL. EDTA will not be dissolved completely in water unless the pH is adjusted to about 8.0.
5. 50× Tris acetate–EDTA (TAE) stock solution: Dissolve 242 g of Tris base (MW = 121.14 g/mol) in approximately 750 mL of deionized water. Carefully add 57.1 mL of acetic acid and 100 mL of 0.5 M EDTA, and add deionized water to make a final volume of 1 L. The pH of this buffer is not adjusted and should be about 8.5.
6. Working solution of TAE buffer (1×): Dilute the stock solution by 50-fold with deionized water. Final solute concentrations are 40 mM Tris acetate and 1 mM EDTA.
7. 0.7% (1%, 2%) agarose gel in 1× TAE buffer: Add 0.7 g (1 g, 2 g) of agarose into 100 mL of 1× TAE buffer, and microwave until agarose is completely melted. Cool the solution to approximately 70–80 °C. Add 5 μL of ethidium bromide into the solution and mix well. Pour 25–30 mL of solution onto an agarose gel rack with appropriate 2-well (for gel purification) or 8-well (for checking PCR products) combs.
8. QIAquick Gel Extraction Kit.
9. NanoDrop 2000c: Used to measure the concentration of DNA.
10. Precision Molecular Mass Standard.
11. pYD1: Obtained from Invitrogen and encodes AGA2, an adhesion subunit of a-agglutinin of a-yeast cells, which enables yeast surface display of a target protein.
12. QIAquick PCR Purification Kit.
13. 1× New England Biolabs buffer (NEBuffer).
14. T4 DNA ligase and 10× T4 DNA ligase buffer.
15. DH5α chemical competent cells: Obtained from Media Preparation Facility (University of Illinois, Urbana, IL) for cloning.

16. 100 mg/mL ampicillin stock solution: Dissolve 1 g of ampicillin powder in 10 mL of ddH_2O and filter-sterilize.
17. Lysogeny broth (LB): Add 10 g of Bacto tryptone, 5 g of yeast extract, 10 g of NaCl into 1 L of ddH_2O, and autoclave.
18. LB-amp agar plates: Autoclave LB agar, and when the solution cools down to 70–80 °C, add 1 mL of 100 mg/mL ampicillin to 1 L of LB agar. Pour 20–25 mL into each Petri dish.
19. QIAprep Miniprep Kit.
20. *S. cerevisiae* EBY100 (*MATa ura3-52 trp1 leu2Δ 1 his3Δ 200 pep4::HIS3 prb1Δ1.6R can1 GAL (pIU211:URA3)*): Obtained from Invitrogen for yeast surface display of protein of interest and plasmid construction through homologous recombination.
21. Zymoprep II yeast plasmid miniprep.
22. 50 mg/mL kanamycin stock solution: Dissolve 1 g of kanamycin powder in 20 mL of ddH_2O and filter-sterilize.
23. Synthetic defined medium with casamino acids (SD-CAA): Dissolve 20 g dextrose, 6.7 g yeast nitrogen base, 10 g casamino acids in 1 L of ddH_2O, and sterilize by autoclaving.
24. SD-CAA agar plates: autoclave SD-CAA agar, and let cool down to 70–80 °C, and add 50 μg/mL kanamycin. Pour 20–25 mL into each Petri dish.
25. OptiKinase and 10× reaction buffer: Obtained from USB and was used to phosphorylate DNA inserts for ligation.
26. Influenza A/Aichi/2/68 viral RNA: A gift from R. Donis (CDC, Atlanta, GA) was used as template to amplify the hemagglutinin gene.
27. SuperScript III One-Step RT-PCR System with Platinum *Taq* High Fidelity.
28. Transcriptor First Strand cDNA Synthesis Kit.
29. DNase I, DNA polymerase I, RNase H, T4 DNA polymerase, and Klenow fragment were obtained from New England Biolabs and stored at −20 °C.
30. QIAEX II Gel Extraction Kit.
31. 10 mg/mL bovine serum albumin (BSA) solution.
32. Shrimp alkaline phosphatase (SAP) and 10× SAP buffer (Fermentas).
33. ElectroMax DH5α competent cells: Obtained from Invitrogen for peptide library construction by electroporation.
34. Influenza virus X31/A/Aichi/68 genomic RNA: Obtained from Charles River Laboratories and used as the template for cDNA synthesis.

2.2 Cell Growth and Transformation

1. Yeast peptone adenine and dextrose (YPAD) medium: Dissolve 10 g of yeast extract, 20 g of peptone, 20 g of dextrose, and 100 mg of adenine hemisulfate in 1 L of ddH_2O, and sterilize by autoclaving.
2. LiAc solutions: To prepare 50 mL of 1 M solution, dissolve 3.3 g lithium acetate (MW = 65.99 g/mol) in ddH_2O, and sterilize by filtration. Prepare 0.1 M solution by mixing 5 mL of the 1.0 M solution with 45 mL sterile ddH_2O.
3. Single-stranded DNA (ssDNA) stock solution: Dissolve 200 mg of deoxyribonucleic acid sodium salt type III from Salmon Testes in 100 mL of Tris–EDTA (TE) buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) by vigorous mixing on a magnetic stirrer. Aliquot the ssDNA in 1 mL and store at −20 °C (*see* **Note 4**).
4. 50% (w/v) polyethylene glycol (PEG): Add 50 g PEG (MW = 3350 g/mol) into 30 mL ddH_2O, heat the solution to ~60 °C with constant stirring to ensure quick dissolution of PEG, and once completely dissolved, adjust volume with ddH_2O to 100 mL, and sterilize by filtration.
5. Transformation mixture: Add 240 μL of 50% PEG, 36 μL of 1 M LiAc, 50 μL of boiled ssDNA, 0.1–10 μg of plasmid DNA, and add sterile ddH_2O to make a final volume of 360 μL. For homologous recombination, 5–20 ng of the digested vector and a 10- to 20-fold molar excess of the insert are used. Transformation mixture should be prepared immediately before use.
6. Gene Pulser II and Pulse Controller Plus: Obtained from Bio-Rad and used to transform plasmids or ligation mixture into *E. coli* through electroporation.
7. 1 M glucose solution: Dissolve 90 g of D-glucose in 400 mL of ddH_2O, adjust to a final volume of 500 mL, and filter-sterilize it.
8. Super Optimal Broth with catabolite repression (SOC) medium: Add 20 g of Bacto tryptone, 5 g of yeast extract, 0.5 g of NaCl, 186.4 mg of KCl into 980 mL of ddH_2O. Adjust the pH to 7.0 with NaOH. Autoclave at 121 °C for 15 min. After the solution cools down to 70–80 °C, add 20 mL of sterile 1 M glucose.

2.3 Expression and Function Analysis of pMHCs on Yeast Cell Surface

1. Yeast peptone galactose (YPG) medium: Dissolve 10 g yeast extract, 20 g peptone and 20 g galactose in 1 L of ddH_2O, and sterilize by autoclaving.
2. Phosphate-buffered saline (PBS).
3. 0.5% BSA in PBS for all antibody staining and washing steps.

4. Antibodies: anti-V5, LB3.1 (American Tissue Culture Collection (ATCC)), and goat anti-mouse (GAM) IgG.
5. Streptavidin-phycoerythrin (SA-PE) conjugate.
6. HA1.7 hybridoma T cells: A gift from J. Bill [38] was used as an indicator cell line for productive DR1-peptide-TCR interactions that produce IL-2 upon receptor engagement.
7. Complete Iscove's Modified Dulbecco's Media (IMDM) for HA1.7 culture: IMDM supplemented with 10% fetal bovine serum and penicillin (10 U/mL)-streptomycin (10 μg/mL).
8. Murine IL-2 enzyme-linked immunosorbent assay (ELISA) kit.

3 Methods

3.1 Vector Construction for Yeast Surface Display of Single-Chain pMHCs

1. Amplify DNA-encoding single-chain DR1α-linker-$HA_{306\text{-}318}$-linker-DR1b using α-5BX and β-3XH as primers (Table 1) and plasmid pJ3/238 as template. Set up the reaction mixture as the following: 10 ng of DNA template, 1× HF Phusion buffer, 50 pmol of forward primer, 50 pmol of reverse primer, 0.25 mM (each) dNTPs, 2 U of Phusion DNA polymerase, and add ddH_2O to make up a final volume of 100 μL. Unless otherwise specified, this recipe was used for all PCR reactions.
2. PCR thermocycler program: an initial denaturation of 2 min at 98 °C, followed by 25–30 cycles of denaturation at 98 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 60 s, and a final 10 min elongation at 72 °C. Unless otherwise specified, this program was used for all PCR reactions with the addition of extension time of 30 s per additional 1 kb long gene.
3. Load the PCR reaction mixture onto 0.7% agarose gels, and perform electrophoresis in 1× TAE buffer at 120 V for 20 min.
4. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit, and check the concentration using NanoDrop, as per the manufacturer's instructions (*see* **Note 5**).
5. Digest 1 μg of the purified PCR product (insert) and 1 μg of plasmid pYD1 (vector) with *Bst*XI and *Xho*I, and purify using the QIAquick PCR Purification Kit.
6. Ligate the digested insert, and vector to generate plasmid pYD1scαHAβ at 16 °C overnight. Set up the ligation mixture as the following: 50 ng of vector, 100 ng of insert, 1× T4 ligase buffer, 1 μL of T4 ligase, and add ddH_2O to make up a final volume of 20 μL.

7. Transform 5 μL of the ligation mixture into chemical competent DH5α cells (*see* **Note 6**), and select transformants on a LB-amp$^+$ agar plate in a 37 °C incubator.
8. Pick a single colony into 3 mL LB medium with 100 μg/mL ampicillin, grow at 37 °C with 250 rpm agitation overnight, isolate the plasmid using the QIAprep Miniprep Kit, as per the manufacturer's instructions, and confirm the sequence of pYD1scαHAβ by DNA sequencing.
9. Introduce a *Not*I restriction site directly upstream of $HA_{306–318}$ (*see* Fig. 1b) to facilitate swapping of different peptides in the single-chain pMHC construct as follows: Perform two PCR reactions using pYD1scαHAβ as template and pYD1For/NotIRev and NotIFor/βRev73-67 as primers (*see* Table 1), respectively. Purify the PCR products using QIAquick PCR Purification Kit, and splice them as described elsewhere [39]. Cotransform the spliced DNA and pYD1scαHAβ digested with *Bst*XI and *Spe*I into *S. cerevisiae* EBY100 to generate pYD1αHAβ. Refer to Subheading 3.3 for detailed yeast transformation protocol. Isolate pYD1αHAβ from a 3 mL yeast culture in SD-CAA using the Zymoprep II yeast plasmid kit.
10. Transform 5 μL of the yeast plasmid into chemical competent DH5α cells (*see* **Note 7**), and isolate the plasmid for DNA sequencing as described in **steps 7** and **8** to confirm the sequence of pYD1αHAβ.
11. Construct plasmid pYD1αSTFβ that contains a segment of stuffer DNA in place of the $HA_{306–318}$ peptide in pYD1αHAβ (*see* **Note 1**): The stuffer DNA (STF) is an unrelated gene (~1 kb) amplified from the phosphite dehydrogenase gene [40] using primers StfFor and StfRev (*see* Table 1). The PCR product was cloned into pYD1αHAβ via NotI and SpeI to generate plasmid pYD1αSTFβ.
12. To establish the correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity (*see* **Note 2**), plasmids pYD1α2467β, pYD1αM4β, pYD1αYAKβ, pYD1αCIIβ, pYD1αPKAβ, pYD1αLβ, and pYD1αMBPβ were constructed in two steps. First, phosphorylate the oligonucleotides by OptiKinase (USB, Cleveland, OH). Set up the phosphorylation reaction as follows: 2 μL of oligonucleotides (2.5 μM), 1× OptiKinase buffer, 2.5 μL of dATP (10 mM), 1 μL of OptiKinase, and add ddH_2O to make a final volume of 25 μL. Incubate the mixture at 37 °C for 2.5–3.5 h and 65 °C for 15 min. Second, generate the double-stranded DNA insert encoding the peptide by hybridization of the forward and reverse oligonucleotides (*see* Table 1). For pYD1αMBPβ, generate the insert using a

self-annealing oligonucleotide MBPNotI. Ligate the resulting insert into pYD1αSTFβ digested by NotI and SpeI.

3.2 Yeast Transformation

1. Inoculate a single colony of *S. cerevisiae* EBY100 into 3 mL of YPAD medium, and grow overnight in a shaker at 30 °C and 250 rpm (*see* **Note 8**).
2. Measure the OD_{600} of the overnight culture, and inoculate the appropriate amount to 50 mL (this is enough for ten transformations, scale up or down proportionally for more or less transformations) of fresh YPAD medium to obtain an OD_{600} of 0.2.
3. Continue to grow the 50 mL of culture for approximately 4–5 h to obtain an OD_{600} of 0.8 (*see* **Note 9**).
4. Wash the cells once with 25 mL of sterile ddH_2O (3380 × *g*, 5 min), once with 1 mL of 0.1 M LiAc (3380 × *g*, 30 s), add 400 μL of 0.1 M LiAc and resuspend (it should result in a total of 500 μL of resuspended cells), and aliquot 50 μL into sterile Eppendorf tubes. Each tube is used for one transformation.
5. In the meantime, boil 1 mL of ssDNA stock solution in a water bath for 5 min and chill immediately on ice.
6. Spin down the 50 μL cells (3380 × *g*, 15 s) and discard the supernatant.
7. Prepare the transformation mixture just before use (*see* Subheading 2.2, **item 5** and *see* **Note 10**).
8. Mix the transformation mixture extensively by vigorous vortexing, incubate at 42 °C for 40 min, spin down the cells (3380 × *g*, 30 s), and remove the transformation mixture carefully.
9. Resuspend the cell pellet in 1 mL ddH_2O gently with pipetting.
10. Plate 50–200 μL of the resuspension on a SD-CAA agar plate, and incubate at 30 °C for 2–3 days until colonies appear.

3.3 Expression Analysis of pMHCs on Yeast Surface

1. Grow *S. cerevisiae* EBY100 clones transformed with different plasmid constructs in 3 mL of SD-CAA at 30 °C until OD_{600} reaches ~5.
2. Wash the yeast cells twice (3380 × *g*, 1 min) with 1 mL YPG medium.
3. To induce AGA2 fusion protein expression, resuspend the cells to an OD_{600} ~1.0 in 3 mL of YPG medium supplemented with 50 μg/mL kanamycin, and incubate in a refrigerated rotary shaker at 20 °C and 250 rpm for ~48 h. The induced yeast cells could be stored at 4 °C up to 4 months for repeated analysis (*see* **Note 11**).

4. For flow cytometric analysis, wash ~2.5 × 10^6 cells in a 96-well V-bottom plate with 180 μL of PBS containing 0.5% BSA, and incubate with the primary antibody (anti-V5) for full-length protein detection and LB3.1 for correctly folded DR1 detection (*see* Fig. 2a) at 4 °C for 1 h at a 100-fold dilution rate.
5. Wash cells once with 180 μL of 0.5% BSA in PBS, and incubate with biotinylated GAM IgG (1:100 dilution) at 4 °C for 1 h.
6. Wash cells once with 180 μL of 0.5% BSA in PBS, and incubate with SA-PE (1:100 dilution) at 4 °C for 30 min.
7. Wash cells three times with 0.5% BSA in PBS to remove unbound SA-PE.
8. Analyze fluorescently labeled yeast cells on a Coulter Epics XL flow cytometer at the Biotechnology Center of University of Illinois at Urbana-Champaign (Urbana, IL) (*see* **Note 12**).

3.4 Stimulation of T Cell Hybridoma and IL-2 Detection

1. Aliquot ~10^6 yeast cells induced with YPG into 300 μL of PBS in a 96-well R-bottom tissue-culture plate. Prepare a triplicate for each pMHC construct.

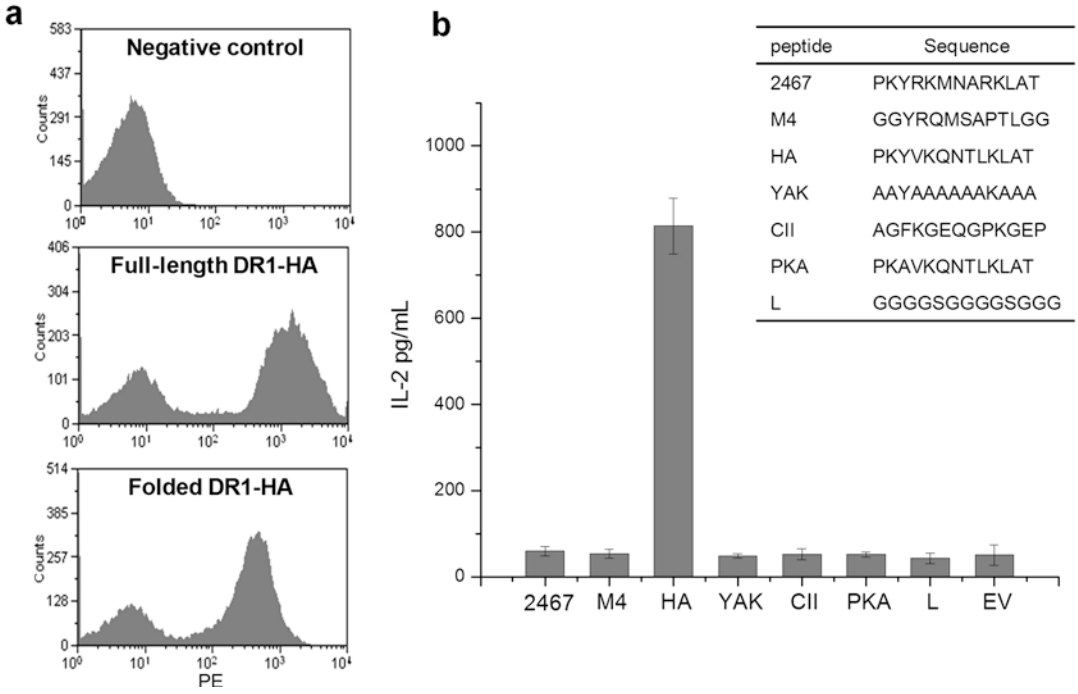


Fig. 2 Expression and function analysis of peptide-DR1 complexes displayed on yeast cell surface. (**a**) Yeast cells transformed with pYD1-scDR1αHAβ were stained with anti-V5 antibody to detect full-length protein expression or with LB3.1 antibody to detect correctly folded complexes. Yeast cells stained only with a secondary antibody were used as a negative control. (**b**) Yeast cells displaying the peptide-DR1 complexes activated HA1.7 hybridoma in an antigen-specific manner. Yeast cells containing the empty plasmid pYD1 (EV) were cultured, induced, and analyzed in the same way as yeast displaying other constructs. Sequences of peptides used in the scDR1αpepβ constructs are shown in the inset

2. Wash once with PBS and resuspend in 300 μL of PBS.
3. Incubate the plate at 4 °C overnight to allow the yeast cells to attach to the surface.
4. Wash away unbound yeast cells three times with 350 μL of PBS.
5. Wash HA1.7 hybridoma cells once with 10 mL of warm complete IMDM medium, and resuspend to a density of 10^5 per 300 μL complete IMDM.
6. Add 300 μL of the HA1.7 hybridoma cell resuspension to each well, and incubate for ~24 h at 37 °C with 5% CO_2.
7. Spin down the cells, and test the supernatant for IL-2 production using the murine IL-2 ELISA kit. IL-2 should only be detected in the supernatant from the wells where the yeast cells displaying functional peptide-MHCs specific for HA1.7 TCR, i.e., DR1-$HA_{306–318}$ (*see* Fig. 2b).

3.5 Peptide Library Construction from Hemagglutinin

1. Amplify DNA-encoding hemagglutinin by One-Step RT-PCR using influenza A/Aichi/2/68 viral ribonucleic acid (RNA) as the template and AichiFor and AichiRev as primers (*see* Table 1). Set up the reaction mixture as follows: 2 μL of RNA, 1× reaction buffer, 0.2 μM of AichiFor, 0.2 μM of AichiRev, 1 μL of SuperScript III, and add ddH_2O to make up a final volume of 50 μL.
2. The thermocycler program: 55 °C for 30 min, an initial denaturation of 2 min at 94 °C, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 68 °C for 2 min, and a final 10 min elongation at 68 °C.
3. Gel-purify the DNA band with correct size using the QIAquick Gel Extraction Kit, and quantify the concentration using NanoDrop, as per the manufacturer's instructions.
4. Fragmentize the purified DNA using DNase I as the following: mix 1 μg of DNA, 5 μL of 500 mM Tris–HCl, pH 7.4, 2.5 μL of 200 mM $MnCl_2$, and ddH_2O to a total volume of 42.5 μL. Incubate the mixture at 15 °C for 15 min. Freshly dilute DNase I with ddH_2O to 0.01 U/μL, and add 7.5 μL to the DNA mixture, and incubate at 15 °C for 1 min. Heat inactivate at 90 °C for 10 min (*see* **Note 13**).
5. Run the DNase I digestion mixture immediately on a 2% agarose gel, and purify the DNA fragments with size between 30 and 60 bp using the QIAEX II Gel Extraction Kit, as per manufacturer's instructions.
6. Blunt-end polish the purified DNA fragments as follows: Mix 0.5 μL of 10 mg/mL BSA, 1× NEBuffer 2, 5 μL of 1 mM dNTP, 100 ng of DNA fragments, 0.1 U of T4 DNA poly-

merase, and add ddH_2O to a final volume of 50 μL. Incubate at 25 °C for 10 min and add 1.25 μL of Klenow fragment (5 U/μL). Incubate at 25 °C for 10 min and 16 °C for 105 min.

7. Gel-purify the blunt-end polished DNA fragments using the QIAEX II Gel Extraction Kit.
8. In the meantime, blunt-end polish pYD1αSTFβ digested by *Spe*I and *Not*I in the same way as described in step 5, gel-purify using QIAquick Gel Extraction Kit, and dephosphorylate using 1 μL of SAP per 1 μg of DNA at 37 °C for 90 min. Heat inactivate at 70 °C for 10 min, and clean up the mixture using the QIAquick PCR Purification Kit.
9. Ligate the DNA fragments from **step 6** into the vector from **step 7** as follows (*see* **Note 14**): Mix 1 μL of T4 DNA ligase buffer, 50 ng of vector, twofold molar excess of DNA fragments, 1.5 μL of PEG 8000, 1 μL of T4 DNA ligase, and add ddH_2O to a final volume of 10 μL. Incubate the ligation mixture at 16 °C for 16–20 h.
10. Clean up the ligation mixture as follows: add *n*-butanol to make a final volume of 500 μL, mix vigorously by vortexing for 30 s, centrifuge at the maximum speed in a benchtop centrifuge for 10 min, immediately remove the supernatant as much as possible, place the tube in a chemical hood for ~ 1 h to dry, and resuspend the ligation mixture in 2 μL of ddH_2O (this is enough for one electroporation).
11. Transform the ligation mixture into ElectroMax DH5α competent cells by electroporation using Gene Pulser II.
12. After electroporation, immediately recover cells in 1 mL of pre-warmed SOC medium at 37 °C with 250 rpm agitation for 1 h.
13. For library creation, perform multiple electroporations, pool the cells into 400 mL LB medium supplemented with 100 μg/mL ampicillin, and grow overnight at 37 °C with 250 rpm agitation. For example, a library of 4.6×10^5 clones was obtained with ten electroporations in our study (*see* **Note 15**).
14. Isolate the plasmids using the QIAprep Miniprep Kit, as per manufacturer's instructions.
15. Perform plasmid transformations into the EBY100 competent yeast cells as described in Subheading 3.3, and pool the cells into 400 mL of SD-CAA medium. With 40 transformations, we obtained a library of 2.1×10^6.
16. After two passages in 400 mL of SD-CAA medium, induce protein expression using YPG as described in Subheading 3.4, and continue with FACS screening.

3.6 Peptide Library Construction from Influenza A Virus

1. Synthesize first-strand cDNA using Transcriptor First Strand cDNA Synthesis Kit with random priming. Set up the reaction as follows: 2 μL of X31/A/Aichi/68 genomic RNA, 2 μL of hexamer, and 9 μL of PCR grade H_2O. Incubate the mixture at 65 °C for 10 min and put on ice immediately. Add 0.5 μL of PCR grade H_2O, 4 μL reaction buffer, 2 μL of dNTP, and 0.5 μL of Transcriptor RTase. Incubate the mixture at 25 °C for 10 min, 55 °C for 30 min, 85 °C for 5 min, and then cool down to 4 °C.
2. Synthesize the second-strand DNA by incubating the cDNA with 10 U of DNA polymerase I, 0.32 U of RNase H, 0.25 mM dNTP, and 1× NEBuffer 2 in a final volume of 40 μL at 15 °C for 90 min.
3. Run the double-stranded DNA on a 1% agarose gel, and purify the bands with different size ranges separately: <500 bp, 500 bp to 1 kb, and >1 kb.
4. Quantify the purified DNA using the Precision Molecular Mass Standard as per the manufacturer's instructions.
5. Fragmentize the DNA with DNase I using the same method as described in Subheading 3.7, **step 3** except for approximately 6, 0.75, and 1.7 U per 1 μg of DNA with size <500 bp, 500 bp to 1 kb, and >1 kb, respectively (*see* **Note 13**).
6. Generate the yeast display library for FACS screening using the same method as described in Subheadings 3.6, **steps 5–16**.

3.7 Yeast Display Library Screening

1. Take an aliquot of the of induced peptide library culture (e.g., 7.4×10^6 yeast cells for the hemagglutinin library and 10^8 for the influenza A virus library), and stain with LB3.1 antibody followed by biotinylated GAM IgG and SA-PE as described in Subheading 3.4.
2. Sort the cells on a Coulter 753 bench FACS sorter (Flow Cytometry Facility, University of Illinois at Urbana-Champaign), and collect ~1.5% of the population with the highest fluorescence (Figs. 3a and 4a) in SD-CAA medium.
3. After protein expression induction in YPG, perform another two rounds of cell sorting (*see* **Note 16**) in the same way as described in **steps 1–2** except that the top 0.5% of the population is collected into SD-CAA medium.
4. Induce protein expression in YPG, and sort the top 0.5% of the population into SD-CAA medium in a 96-well plate with no more than one cell in a well.
5. For hemagglutinin-derived peptide library, randomly pick ten of these cells, and examine surface protein display level and their ability to stimulate T cell hybridoma HA1.7 as described in Subheadings 3.4 and 3.5 (Fig. 3b, c). To determine the

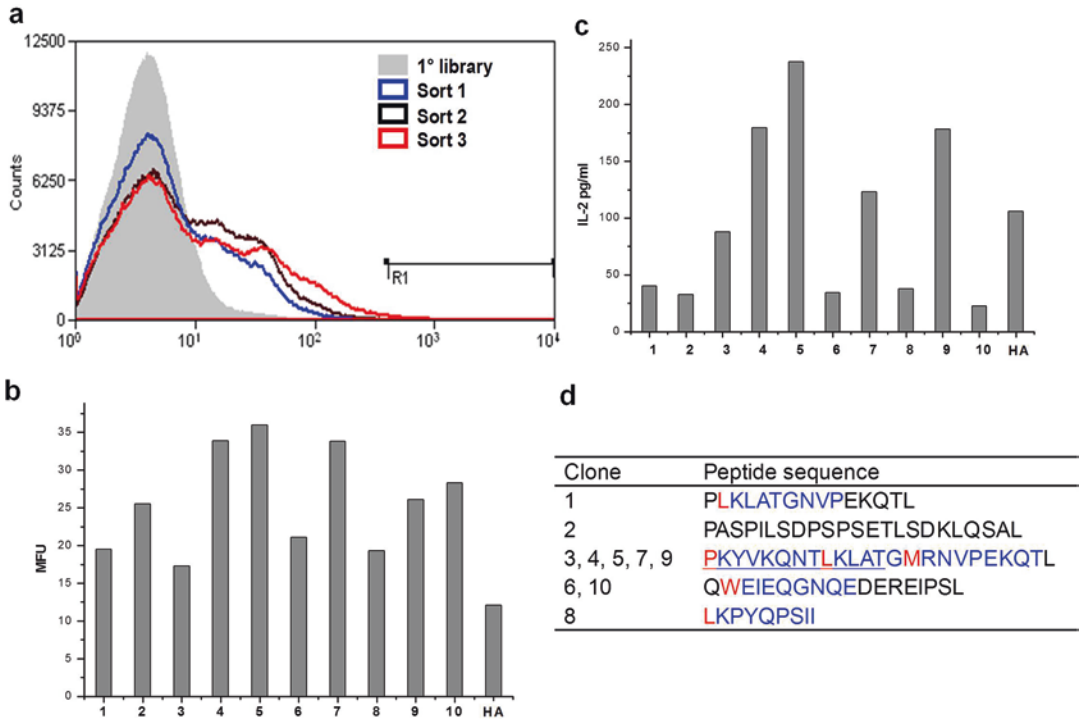


Fig. 3 Direct epitope identification from a single antigenic protein with known sequence—hemagglutinin. (**a**) FACS enrichment of potential good binders from the hemagglutinin-derived peptide library. LB3.1 antibody was used to stain cells as a measurement of surface expression level of correctly folded peptide-DR1 complexes. Surface expression (**b**), T cell hybridoma activation (**c**), and DNA sequence (**d**) analysis of ten clones randomly picked from the library after three rounds of cell sorting. The predicted binding 9-residue peptides are shown in blue with red letters corresponding to the amino acid residue at position P1

sequence of the peptide insert, perform DNA sequencing analysis.

6. For influenza A virus-derived peptide library, sort four 96-well plates of cells from the top 0.5% of the population, and analyze them in the same way as described in **step 5** (*see* Fig. 4b, c).
7. Align peptide sequences from the active clones to identify epitope sequences using ClustalW (http://embnet.vital-it.ch/software/ClustalW.html) (*see* Figs. 3d and 4d).

4 Notes

1. To eliminate the possibility that the epitope $HA_{306–318}$ identified from the peptide library is derived from the undigested pYD1αHAβ, it is important to use plasmid pYD1αSTFβ that contains a segment of unrelated stuffer DNA in place of the $HA_{306–318}$ peptide as the backbone for blunt-end ligation. In

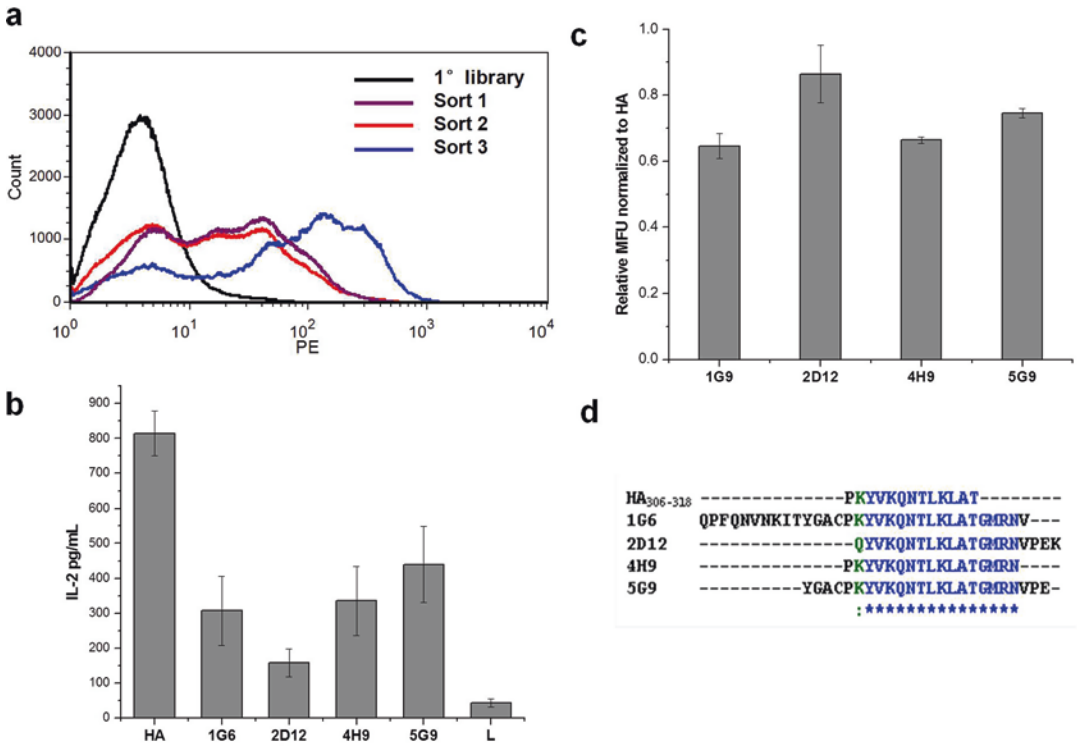


Fig. 4 Direct epitope identification from a pathogen with unknown sequence—influenza A virus. (**a**) FACS enrichment of potential DR1-binding peptides using LB3.1. (**b**) T cell activation analysis of the four clones identified from the enriched peptide library. (**c**) The four active clones showed comparable surface expression levels as yeast displaying scDR1αHAβ. The relative mean fluorescence unit (MFU) was normalized to HA for direct comparison. (**d**) The peptide sequences from the four clones were aligned with the $HA_{306–318}$ epitope sequence. Asterisk, fully conserved residues, conservation of strong groups

addition, the 1 kb stuffer DNA enables clear separation of digested plasmid by DNA electrophoresis.

2. A positive correlation between the yeast surface expression level of the peptide-DR1 complexes and the peptide-binding affinity was established [37]. Therefore, the yeast surface expression level of the single-chain DR1-peptide protein, as measured by DR-specific antibody (LB3.1) staining, can be used as a proxy screening variable for DR1-binding peptides, significantly reducing the number of clones required in the function screening assay. This stabilizing effect of a binding peptide was also observed for DR4 [37] and DR2 [41].
3. The ability of yeast cells displaying DR1-$HA_{306–318}$ to activate T hybridoma cell HA1.7 greatly simplified the functional screening for epitope identification. This could be generally applicable to MHCII alleles, such as DR4 [37]. However, due to the polygenic and polymorphic nature of MHC proteins and the diversity of the binding peptides, it is difficult to establish that

yeast cells could display all or most of pMHCs in a functional form to engage specific TCRs, such as low-affinity self-reactive ones (e.g., DR2-MBP85-99-Ob.1A12 [41]).

4. It is not desirable to freeze-thaw the carrier ssDNA frequently, and usually it is discarded after three or four times of thawing.
5. When the DNA concentration is low (below 20 ng/μL), run 3–5 μL on a 0.7% agarose gel, and use the Precision Molecular Mass Standard for a more accurate estimation according to the manufacturer's instructions.
6. DH5α was used for DNA cloning in our experiments. However, any *E. coli* strain with *recA* and *endA*, such as Top10 and JM109, can also be used.
7. The number of DH5α transformants could vary from a few to several thousands. This is mainly due to the low quality of the isolated yeast plasmids. Sometimes, no colony was even obtained. In such cases, repeat the DH5α transformation.
8. Yeast competent cells need to be freshly prepared each time.
9. Normally, the doubling time for a *S. cerevisiae* laboratory strain is approximately 2 h.
10. It is important to follow the order by which the transformation mixture components are added, especially for 50% PEG, which shield the competent cells from the toxic 1 M LiAc.
11. The storage time might vary depending on the stability of the target protein displayed on yeast cell surface. For the pMHCs in our study, no significant degradation was observed within 4 months.
12. To compare surface expression levels of different pMHCs, it is important to induce the protein expression and perform the antibody staining in a single experiment run to minimize variation.
13. Since DNase I has very high activity, it is necessary to test and optimize the digestion conditions for different target DNA preparations and for different lots of DNase I.
14. The difficulties in cloning small pieces of DNA include (a) optimization of the DNase I digestion step to have majority of the digested DNA fragments with a desired size range; (b) intramolecular circularization, which results in a high ligation background; and (c) occurrence of multiple inserts, which results in splicing peptides that are not present in the natural peptide repertoire. Accordingly, when constructing the peptide library, it is important to (a) optimize the DNase I concentration and the digestion time; (b) dephosphorylate the vector and include 15% PEG in the ligation mixture to promote intermolecular ligation over intramolecular circularization [42]; and (c) optimize the insert to vector molar ratio.

15. It is important to sample the diversity of the primary peptide library created in *E. coli*. To do so, an aliquot of the library is selected on a LB-amp agar plate to obtain individual colonies (this also allows an estimate of the library size), and 20 clones are randomly picked, sequenced, and mapped to the viral genome. There should be at least one peptide derived from each of the eight pieces of influenza viral genomic RNA displayed in either sense or antisense orientations, indicating a good representation of the whole viral genome.
16. The number of sorting cycles varies depending on how much improvement of the expression the enriched library has over the previous round. For example, there was a steady increase in the fluorescence intensity over the sorting rounds in Fig. 3a, so a total of four rounds of enrichment (the fourth round is indicated by region R1) were performed.

References

1. Germain RN (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell 76:287–299
2. Wen F, Rubin-Pitel SB, Zhao H (2009) Engineering of therapeutic proteins. In: Park SJ, Cochran JR (eds) Protein engineering and design. Tayler & Francis Group, Boca Raton, FL
3. Park H, Li Z, Yang XO et al (2005) A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat Immunol 6:1133–1141
4. Altman JD, Moss PA, Goulder PJ et al (1996) Phenotypic analysis of antigen-specific T lymphocytes. Science 274:94–96
5. Klein L, Kyewski B (2000) "Promiscuous" expression of tissue antigens in the thymus: a key to T-cell tolerance and autoimmunity? J Mol Med 78:483–494
6. Huang CH, Peng S, He L et al (2005) Cancer immunotherapy using a DNA vaccine encoding a single-chain trimer of MHC class I linked to an HPV-16 E6 immunodominant CTL epitope. Gene Ther 12:1180–1186
7. Coulie PG, Karanikas V, Lurquin C et al (2002) Cytolytic T-cell responses of cancer patients vaccinated with a MAGE antigen. Immunol Rev 188:33–42
8. Hill BD, Zak A, Khera E et al (2018) Engineering virus-like particles for antigen and drug delivery. Curr Protein Pept Sci 19:112–127
9. Jenkins MK, Khoruts A, Ingulli E et al (2001) In vivo activation of antigen-specific CD4 T cells. Annu Rev Immunol 19:23–45
10. Robbins PF, Morgan RA, Feldman SA et al (2011) Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. J Clin Oncol 29:917–924
11. Morgan RA, Dudley ME, Wunderlich JR et al (2006) Cancer regression in patients after transfer of genetically engineered lymphocytes. Science 314:126–129
12. Phan GQ, Yang JC, Sherry RM et al (2003) Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. Proc Natl Acad Sci U S A 100:8372–8377
13. Matsushita H, Vesely MD, Koboldt DC et al (2012) Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature 482:400–404
14. (2017) The problem with neoantigen prediction. Nat Biotechnol 35:97
15. Schumacher TN, Schreiber RD (2015) Neoantigens in cancer immunotherapy. Science 348:69–74
16. Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC class I binding peptides using profile motifs. Hum Immunol 63:701–709
17. Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. BMC Bioinformatics 8:238
18. Hattotuwagama CK, Doytchinova IA, Flower DR (2007) Toward the prediction of class I and II mouse major histocompatibility

complex-peptide-binding affinity: in silico bioinformatic step-by-step guide using quantitative structure-activity relationships. Methods Mol Biol 409:227–245

19. Huang M, Huang W, Wen F et al (2017) Efficient estimation of binding free energies between peptides and an MHC class II molecule using coarse-grained molecular dynamics simulations with a weighted histogram analysis method. J Comput Chem 38:2007–2019
20. Nielsen M, Lundegaard C, Worning P et al (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12:1007–1017
21. Andreatta M, Karosiene E, Rasmussen M et al (2015) Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. Immunogenetics 67:641–650
22. Luo H, Ye H, Ng HW et al (2015) Machine learning methods for predicting HLA-peptide binding activity. Bioinform Biol Insights 9:21–29
23. Nielsen M, Lund O, Buus S et al (2010) MHC class II epitope predictive algorithms. Immunology 130:319–328
24. Wang P, Sidney J, Dow C et al (2008) A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 4:e1000048
25. Jensen KK, Andreatta M, Marcatili P et al (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology 154:394–406
26. Vita R, Overton JA, Greenbaum JA et al (2015) The immune epitope database (IEDB) 3.0. Nucleic Acids Res 43:405–412
27. Lin HH, Zhang GL, Tongchusak S et al (2008) Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. BMC Bioinformatics 9(Suppl 12):S22
28. Henderson RA, Cox AL, Sakaguchi K et al (1993) Direct identification of an endogenous peptide recognized by multiple HLA-A2.1-specific cytotoxic T cells. Proc Natl Acad Sci U S A 90:10275–10279
29. Richards KA, Chaves FA, Sant AJ (2009) Infection of HLA-DR1 transgenic mice with a human isolate of influenza a virus (H1N1) primes a diverse CD4 T-cell repertoire that includes CD4 T cells with heterosubtypic cross-reactivity to avian (H5N1) influenza virus. J Virol 83:6566–6577
30. Babon JA, Cruz J, Orphin L et al (2009) Genome-wide screening of human T-cell epitopes in influenza A virus reveals a broad spectrum of CD4+ T-cell responses to internal proteins, hemagglutinins, and neuraminidases. Hum Immunol 70:711–721
31. Novak EJ, Liu AW, Gebe JA et al (2001) Tetramer-guided epitope mapping: rapid identification and characterization of immunodominant CD4+ T cell epitopes from complex antigens. J Immunol 166:6665–6670
32. Smith MR, Tolbert SV, Wen F (2018) Protein-scaffold directed nanoscale assembly of T cell ligands: artificial antigen presentation with defined valency, density, and ratio. ACS Synth Biol 7:1629–1639
33. Sospedra M, Pinilla C, Martin R (2003) Use of combinatorial peptide libraries for T-cell epitope mapping. Methods 29:236–247
34. Hiemstra HS, Duinkerken G, Benckhuijsen WE et al (1997) The identification of CD4+ T cell epitopes with dedicated synthetic peptide libraries. Proc Natl Acad Sci U S A 94: 10313–10318
35. Wang RF, Wang X, Atwood AC et al (1999) Cloning genes encoding MHC class II-restricted antigens: mutated CDC27 as a tumor antigen. Science 284:1351–1354
36. Koelle DM (2003) Expression cloning for the discovery of viral antigens and epitopes recognized by T cells. Methods 29:213–226
37. Wen F, Esteban O, Zhao H (2008) Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. J Immunol Methods 336:37–44
38. Boen E, Crownover AR, McIlhaney M et al (2000) Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. J Immunol 165:2040–2047
39. Horton RM, Cai ZL, Ho SN et al (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. BioTechniques 8:528–535
40. Woodyer R, van der Donk WA, Zhao H (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry 42:11604–11614
41. Wen F, Sethi DK, Wucherpfennig KW et al (2011) Cell surface display of functional human MHC class II proteins: yeast display versus insect cell display. Protein Eng Des Sel 24:701–709
42. Upcroft P, Healey A (1987) Rapid and efficient method for cloning of blunt-ended DNA fragments. Gene 51:69–75
43. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15:553–557



Chapter 14

Isolation of Major Histocompatibility Complex (MHC)-Associated Peptides by Immunoaffinity Purification

Sarah A. Penny and Stacy A. Malaker

Abstract

MHC molecules on the surface of cells are responsible for the presentation of antigenic peptides to CD8+ and CD4+ T cells. Downstream analysis of these peptides can offer insight into various disease processes with an immune component, as in autoimmune diseases and cancer. A critical step in studying MHC-associated peptides is their isolation from MHC molecules on the surface of cells. In this chapter, we detail anti-MHC antibody conjugation to beads, immunoaffinity purification of the MHC molecules, and peptide elution. This method can be used for analysis of unmodified and/or posttranslationally modified peptides by mass spectrometry.

Key words MHC-associated peptide isolation, Immunoaffinity purification, Peptide elution

1 Introduction

Cells are equipped with efficient systems of displaying their health status to the immune system, namely, the major histocompatibility complex (MHC) processing pathways. The MHC class I pathway presents peptides from endogenous proteins to CD8+ T cells, whereas the MHC class II processing pathway presents peptides derived from exogenous or membrane-bound proteins to CD4+ T cells [1, 2]. Activated CD8+ and CD4+ T cells can initiate a strong adaptive immune response. As such, MHC-peptide complexes are implicated in almost every autoimmune disorder (diabetes [3], psoriasis [4], arthritis [5]), as well as several other diseases such as narcolepsy [6] and cancer [7]. It follows, then, that a considerable amount of research is devoted to analysis of MHC-associated peptides, most often by sequencing with mass spectrometry [8, 9].

Here, we present a robust method for isolating MHC-associated peptides from cells or tissues [10, 11]. The sample is first lysed, then proteins are purified from other cell components, and MHC moleccules are immunoaffinity purified. During lysis, the zwitterionic detergent present in the lysis buffer (3-((3-cholamidopropyl)

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_14, © Springer Science+Business Media, LLC, part of Springer Nature 2019

dimethylammonio)-1-propanesulfonate (CHAPS)) disrupts the cells without disrupting the MHC-peptide complexes. Thus, the complexes are maintained for antibody purification, enabling subsequent acid elution of the peptide. Protein purification is achieved using centrifugation, and the batch immunopurification is performed using antibody-bound N-hydroxysuccinimide (NHS) ester-activated Sepharose beads. The antibody is bound to the Sepharose beads by reacting primary amines on the antibody to the NHS ester-activated resin in slightly alkaline conditions (pH 7.2–8.5), yielding stable amide bonds. This non-specific reaction means that this method can be utilized with any antibody and also that the antibody structure is preserved. After the antibody binds the MHC, the beads are subjected to a series of washes with different salt concentrations to remove all impurities. Subsequently, the peptides are acid-eluted from the MHC-peptide complexes and separated from the MHC class I heavy chain and β2-microglobulin by filtration. The immunopurification of the MHC of choice prior to acid elution makes this method highly specific. In our hands, this protocol has allowed for robust peptide analysis from several cancers, cell lines, and human tissues [12, 13].

2 Materials

All buffers must be made in ultrapure water. Great care should be taken to avoid introducing contaminants from skin, plastic, or hair. Reagents can be obtained from Sigma-Aldrich unless otherwise specified. Buffers are kept at 4 °C, unless otherwise specified.

2.1 Reagents and Buffers

1. Phosphate-buffered saline (PBS) (sterile).
2. 0.2 M acetic acid.
3. 10% acetic acid.
4. 1 M HCl.
5. Tris–HCl buffer 1: 0.1 M Tris–HCl, pH 9, 500 mL.
6. Tris–HCl buffer 2: 20 mM Tris–HCl, 150 mM NaCl, pH 8, 500 mL.
7. Tris–HCl buffer 3: 20 mM Tris–HCl, 1 M NaCl, pH 8, 500 mL.
8. Tris–HCl buffer 4: 20 mM Tris–HCl, pH 8, 500 mL.
9. Acetate buffer: 0.1 M acetate, 0.5 M NaCl, pH 5, 500 mL.
10. Lysis buffer: 20 mM Tris–HCl, 150 mM NaCl, 1% CHAPS, pH 8, 500 mL.
11. Aprotinin: 5 mg/mL in lysis buffer (*see* **Note 1**).
12. Leupeptin: 10 mg/mL stock solution in lysis buffer.

13. Pepstatin A: 5 mg/mL stock solution in methanol.
14. 500 mM ethylenediaminetetraacetic acid (EDTA).
15. 20% sodium azide in ultrapure water.
16. Phenylmethylsulfonyl fluoride (PMSF): 17.4 mg/mL stock solution in isopropanol.
17. Phosphatase inhibitor cocktail II (contains sodium orthovanadate, sodium molybdate, sodium tartrate, and imidazole).
18. Phosphatase inhibitor cocktail III (contains cantharidin, (−)-p-Bromolevamisole oxalate, and calyculin A).

2.2 Antibodies for Isolation of Human MHC (See Note 2)

Due to the highly polymorphic nature of the human leukocyte antigens (HLA) that make up the MHC molecules, careful antibody selection is of vital importance to ensure that the HLA for the required investigation is targeted. Antibodies are commercially available, but they may also be produced in-house using hybridomas (*see* **Note 3**) (Tables 1 and 2) [14].

Table 1
Human leukocyte antigen (HLA) class I antibodies

Target	Clone/antibody
Pan class I	W6/32
HLA-A2	BB7.2
HLA-A3	GAPA3
HLA-A11/A24	A11.1M
HLA-B7	BB7.1
HLA-B7/B27	B27M1

Table 2
Human leukocyte antigen (HLA) class II antibodies

Target	Clone/antibody
Pan class II	IVA12
HLA-DP	HI43
HLA-DQ	IVD12
HLA-DR	LB3.1

2.3 Equipment

1. Fast performance liquid chromatography (FPLC) instrument.
2. NHS-Sepharose beads.
3. Tube revolver/rotator.
4. Ultracentrifuge.
5. NanoDrop spectrophotometer.
6. Tissue homogenizer that allows homogenization on ice (such as Qiagen TissueRuptor, Fisher Scientific) (*see* **Note 4**).
7. Protein low-bind microcentrifuge tubes.
8. 5000 molecular weight cutoff (MWCO) filters.
9. Refrigerated microcentrifuge.
10. Ultracentrifuge tubes.

3 Methods

3.1 Preparation of Samples for MHC Immunopurification

MHC-associated peptides can be isolated from cell lines or tissue samples. Follow Subheading 3.1.1 or 3.1.2 accordingly, based on the starting material.

3.1.1 Cell Lines

1. Culture the cell lines to be investigated according to American Type Culture Collection (ATCC) manuals.
2. Pellet cells to remove culture medium.
3. Determine cell number using a hemocytometer.
4. Resuspend in cold PBS buffer (10 mL/1 × 10^8 cells).
5. Pellet cells at 4 °C, and store at −80 °C until required for MHC immunopurification (*see* **Note 5**).

3.1.2 Tissue

1. After tissue is surgically removed, it should be transferred to the laboratory in sterile tubes as quickly as possible.
2. Weigh the tissue and note the weight.
3. Freeze the tissue to −80 °C as soon as possible, until required for MHC immunopurification (*see* **Note 5**).

3.2 Bead Preparation for Immunopurification

1. Determine the amount of antibody needed for extraction: 5 mg of antibodies are used for 1 × 10^9 cells or 1 g of tissue.
2. Calculate the volume of NHS-Sepharose beads necessary: 100 μL of beads per 1 mg of antibody.
3. On ice, use a cut-off pipette tip to move the NHS-Sepharose beads into a centrifuge tube.
4. Add 10 mL of cold, sterile PBS to the beads. Centrifuge at 4 °C, 500 × *g*, for 2 min. Remove supernatant, being careful to not disrupt the bead pellet. Repeat once (*see* **Note 6**).

5. Add the required antibody to the NHS-Sepharose beads (*see* **Note 7**).
6. Rotate at 4 °C overnight.
7. The following day, pellet the beads (4 °C, 500 × *g*, for 2 min) and aspirate the supernatant into another tube.
8. Remove a small sample of this supernatant and acidify to measure the optical density (OD) A_{280} to determine how much antibody is bound to the beads (*see* **Note 8**).
9. To the beads, add 10 mL of Tris–HCl buffer 1 for 1 h, rotating at 4 °C (*see* **Note 6**).
10. Pellet the beads (4 °C, 500 × *g*, for 2 min) and remove the supernatant.
11. Wash the beads in 10 mL of acetate buffer and pellet the beads. Remove the supernatant and discard.
12. Wash the beads in 10 mL of Tris–HCl buffer 1 and pellet the beads. Remove the supernatant and discard.
13. Repeat **steps 10** and **11**.
14. Resuspend the beads in Tris–HCl buffer 2 to the equivalent of 10 mg of antibody/mL buffer.

3.3 Cell Lysis

1. Add protease/phosphatase inhibitors to the lysis buffer. Use 10 mL of lysis buffer per 1×10^9 cell equivalents or 1 g of tissue (*see* **Note 1**). Add 1:1000 (volume/volume) of aprotinin stock solution, 1:1000 of leupeptin stock solution, 1:500 of pepstatin A stock solution, 1:100 of EDTA stock solution, and 1:500 of sodium azide stock solution. At the last second, add 1:100 of PMSF stock solution and protease inhibitor cocktails II and III.
2. Remove cell pellets or tissue from the freezer onto ice; do not thaw. Immediately add cold lysis buffer and disrupt the pellet. If tissue is being used, it must be physically disrupted, after the lysis buffer has been added, without heating, using a tissue homogenizer (*see* **Note 9**).
3. Rotate gently at 4 °C for 1–2 h.
4. During this time, pre-wash and pre-chill the ultracentrifuge tubes, using ultrapure water.
5. Transfer the lysate back to ice and transfer into labeled glass ultracentrifuge tubes.
6. Balance the tubes with fresh lysis buffer to within 0.05 g (*see* **Note 10**).
7. Ultracentrifuge for 1 h at 100,000 × *g* at 4 °C.

3.4 Immunoaffinity Purification of MHC and Elution of Associated Peptides

1. Transfer the supernatant from the pelleted lysate into fresh centrifuge tubes.
2. Add antibody-conjugated beads to the supernatant such that approximately 5 mg of antibody is used per 1×10^9 cell equivalents or 1 g tissue.
3. Use Parafilm to secure the lid on the centrifuge tubes and prevent any leakage.
4. Rotate at 4 °C overnight.
5. The following day, pellet the beads by centrifuging at 4 °C, 500 × *g*, for 2 min.
6. Remove the supernatant and put into a fresh centrifuge tube and store at −80 °C (*see* **Note 11**).
7. Wash the beads once in 10 mL of lysis buffer, and pellet by centrifuging at 4 °C, 500 × *g*, for 2 min.
8. Remove supernatant, and transfer the beads to a low-bind microcentrifuge tube.
9. Wash the beads in 1 mL of Tris–HCl buffer 2. Pellet by centrifuging at 4 °C, 500 × *g*, for 2 min, and discard the supernatant. Repeat.
10. Wash the beads in 1 mL of Tris–HCl buffer 3. Pellet by centrifuging at 4 °C, 500 × *g*, for 2 min, and discard the supernatant. Repeat.
11. Wash the beads in 1 mL of Tris–HCl buffer 4. Pellet by centrifuging at 4 °C, 500 × *g*, for 2 min, and discard the supernatant. Repeat twice for a total of three washes (*see* **Note 12**).
12. During the washes, pre-wet the 5000 MWCO membrane with 500 μL of Tris-HCl buffer 4, and centrifuge at 4 °C, 5000 × *g*, for 20 min.
13. Remove all liquid from the filter unit.
14. Transfer all of the beads into the top of the filter unit and perform a short spin (5000 × *g* for 10 s) to remove excess supernatant.
15. Centrifuge for 30 min at 5000 × *g* to remove all excess buffer.
16. Transfer the filter unit into a new low-bind microcentrifuge tube.
17. To elute the MHC-associated peptides, add 200 μL of 10% acetic acid, and spin for 99 min at 5000 × *g*.
18. Add another 200 μL of 10% acetic acid and spin for 99 min 5000 × *g*.
19. Remove the filter insert and discard.

20. Place the lid onto the microcentrifuge tube, wrap with Parafilm, and store at −80 °C until required for downstream analysis, such as peptide identification by mass spectrometry.

4 Notes

1. Due to the low stability of some of the protease/phosphatase inhibitors in water, these are added from stock solution to the CHAPS lysis buffer just prior to use, as detailed in Table 3.
2. Selection of antibody is dependent on the goal(s) of the project. If an investigator is interested in peptides presented by all class I HLA molecules, a pan class I antibody should be used. If the investigator is interested in peptides presented by a specific HLA allele, then an antibody specific for that allele should be chosen (i.e., A2, B7, DQ, etc.). The same goes for class II HLA molecules.
3. The relevant antibodies may be commercially available (*see* Tables 1 and 2), but due to the large quantities needed in MHC-peptide isolation, it may be more cost-effective to produce antibodies in-house using hybridoma technologies. Antibodies can then be purified from the hybridoma supernatant using protein A Sepharose columns [14].
4. The tissue homogenization produces heat and therefore to stop degradation of proteins it needs to be performed on ice.

Table 3
Protease inhibitor cocktail components

Protease inhibitor	Stock solution concentration	Ratio in lysis buffer (v/v)
5 μg/mL aprotinin	5 mg/mL in lysis buffer	1:1000
10 μg/mL leupeptin	10 mg/mL in lysis buffer	1:1000
10 μg/mL pepstatin A	5 mg/mL in methanol	1:500
5 mM EDTA	500 mM EDTA	1:100
0.04% sodium azide	20% in ultrapure water	1:500
1 mM PMSF (add immediately before use)	17.4 mg/mL in isopropanol	1:100
Phosphatase inhibitor cocktail II and III (add immediately before use)		1:100

5. No cryopreservation is needed, as the cells/tissue will be lysed. This method cannot be used with formalin-fixed tissue.
6. As the antibody is bound to the Sepharose beads by reacting primary amines on the antibody to the NHS ester-activated resin in a non-specific manner, primary amine buffers, such as Tris-buffered saline (TBS), cannot be used. PBS is used instead. Tris buffer is later used to quench the reaction at the end of the conjugation procedure and ensure that the beads will not react with other amine groups from the cell lysate.
7. The antibody must be in suspension at a concentration of around 1–2 mg/mL—concentrated enough to conjugate to the beads, but not so concentrated that it falls out of solution.
8. After the conjugation is completed, the concentration of antibody should have reduced to around 0.1 mg/mL. Unbound antibody is removed through a series of wash steps.
9. To prevent degradation of the protein complexes, all steps must be performed at 4 °C. Homogenization is a heat-producing process; so to minimize heating, this must be performed with the sample on ice, for a minimal time.
10. For safety reasons, ultracentrifuge rotors must always be well balanced with tubes of equal weight opposite one another.
11. Isolation of peptides from multiple HLA alleles can be accomplished by sequential immunoaffinity purifications. This applies to different alleles within the same type of MHC (I or II) and also isolation of both classes of MHC.
12. It is important for subsequent analysis using mass spectrometry that all traces of salt are removed.

Acknowledgments

The authors would like to acknowledge Mark Cobbold (MGH) and Donald F. Hunt (U Virginia) for their technical and financial support. S.A.M. is currently funded by an NIH F32 postdoctoral fellowship.

References

1. Blum JS, Wearsch PA, Cresswell P (2013) Pathways of antigen processing. Annu Rev Immunol 31:443–473
2. Neefjes J, Jongsma ML, Paul P, Bakke O (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol 11:823–836
3. Tisch R, McDevitt H (1996) Insulin-dependent diabetes mellitus. Cell 85:291–297
4. Nair RP, Stuart PE, Nistor I, Hiremagalore R, Chia NVC, Jenisch S, Weichenthal M, Abecasis GR, Lim HW, Christophers E, Voorhees JJ, Elder JT (2006) Am J Hum Genet 78(5): 827–851

5. Nepom GT, Erlich H (1991) MHC class II molecules and autoimmunity. Ann Rev Immunol 9:493–525
6. Bernard-Valnet R, Yshii L, Queriault C, Nguyen XH, Arthaud S, Rodrigues M, Canivet A, Morel AL, Matthys A, Bauer J, Pignolet B, Dauvilliers Y, Peyron C, Liblau RS (2016) CD8 T cell-mediated killing of orexinergic neurons induces a narcolepsy-like phenotype in mice. Proc Natl Acas Sci USA 113(39): 10956–10961
7. Dunn GP, Old LJ, Scheiber RD (2004) The immunobiology of cancer immunosurveillance and immunoediting. Immunity 21(2):137–148
8. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir R, Cox AL, Appella E, Engelhard VH (1992) Characterization of peptides bound to the class-I MHC molecule HLA-A2.1 by mass spectrometry. Science 255(5049):1261–1263
9. Chicz RM, Urban RG, Gorga JC, Vignali DAA, Lane WS, Strominger JL (1993) Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. J Exp Med 178(1):27–47
10. Zarling AL, Ficarro SB, White FM, Shabanowitz J, Hunt DF, Engelhard VH (2000) Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo. J Exp Med 192(12):1755–1762
11. Zarling AL, Polefrone JM, Evans AM, Mikesh LM, Shabanowitz J, Lewis ST, Engelhard VH, Hunt DF (2006) Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc Acad Natl Sci USA 103(40):14889–14894
12. Malaker SA, Penny SA, Steadman LG, Myers PT, Loke JC, Raghavan M, Bai DL, Shabanowitz J, Hunt DF, Cobbold M (2017) Identification of glycopeptides as posttranslationally modified neoantigens in leukemia. Cancer Immunol Res 5(5):376–384
13. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Cummings KL, Penny SA, Turner JE, Cottine J, Abelin JG, Malaker SA et al (2013) MHC Class I–Associated Phosphopeptides Are the Targets of Memory-like Immunity in Leukemia. Sci Transl Med 5(203):203ra125
14. Winzeler A, Wang JT (2013) Culturing hybridoma cell lines for monoclonal antibody production. Cold Spring Harb Protoc 2013(7): 640–642



Chapter 15

Enhanced Mass Spectrometry Detection of MHC Peptides

Rui Chen and Jianjun Li

Abstract

Mass spectrometry (MS)-based immunopeptidomics (MHC peptides) is a promising approach for neoantigen discovery through searching MS data against patient-specific protein databases built from exome or transcriptome sequences. MS analysis enables mapping of posttranslational modifications that cannot be predicted from genome sequencing alone but can be more efficient in triggering immune responses. Although MS-based immunopeptidomics has demonstrated its potential in discovering neoantigens as vaccine candidates, the problems such as efficient isolation and sensitive detection (identification) of MHC peptides still hinder the application of its clinical application. Herein, we describe a method based on chemical derivatization and LC-MS/MS for the enhanced detection of MHC peptides.

Key words MHC peptide, Mass spectrometry, Immunopeptidomics, Mild acid elution, Immunoprecipitation, Chemical derivatization, Proteogenomics

1 Introduction

Neoantigen-based therapeutic vaccines have the potential to impact tumor eradication and cancer patient survival. Mass spectrometry (MS)-based immunopeptidomics has the capacity to identify tumor-associated epitopes. The strength of this approach is in identifying mutation-bearing major histocompatibility complex (MHC)-binding peptides, which have potential therapeutic cancer vaccines [1, 2]. While promising, this approach presents several challenges, including the isolation of low-abundance peptides from various types of samples and identification of the sequence of those peptides bearing mutations. The analysis of immunopeptidome using mass spectrometry requires a large number of cells with an expression rate of $\sim 2 \times 10^5$ MHC molecules per cell, e.g., $\sim 5 \times 10^8$ cells. Thus it is vital to ensure the specific isolation of sufficient numbers of MHC peptides from cells. Mild acid elution (MAE) and immunoprecipitation (IP) are well-established approaches for isolating MHC-bound peptides (Fig. 1), although both methods still require a large quantity of biomass.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_15, © Springer Science+Business Media, LLC, part of Springer Nature 2019

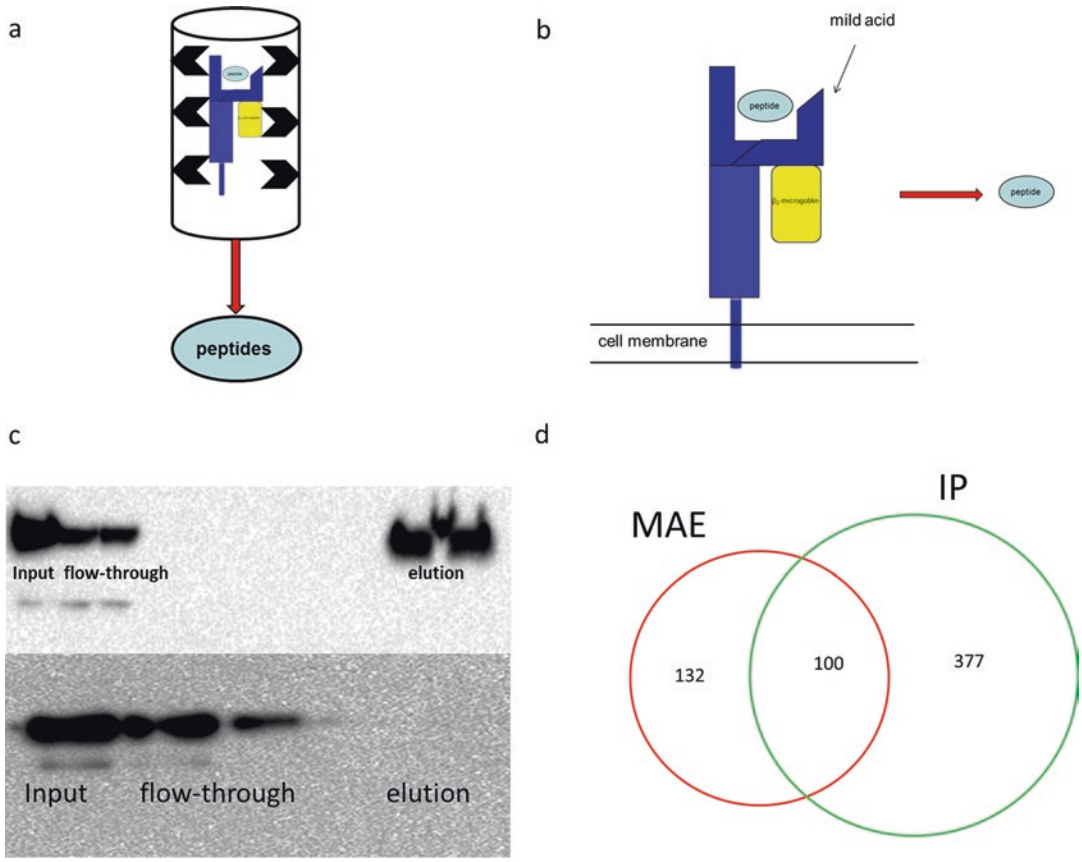


Fig. 1 Isolation of MHC I peptides from cell lines with immunoprecipitation and mild acid elution. (**a**) Scheme for immunoprecipitation of MHC I with W6/32 antibody; (**b**) scheme for mild acid elution of MHC I peptides from cell surface; (**c**) Western blotting of immunoprecipitation from native cell lysate (top) and lysate from cells after mild acid elution (bottom); (**d**) overlap between MHC I peptides identified by LC-MS/MS from IP and MAE

MAE is a rapid approach that elutes the MHC peptides by dissociating the MHC complex with acidic buffer. However, it lacks selectivity, and an estimated 40–60% of the identified peptides can be contaminants with low binding affinity to human leukocyte antigen (HLA) [3]. In comparison, the IP strategy has the advantage of high specificity and flexibility, which can be used for the isolation of both class I and class II MHC peptides from a range of biological sources such as cultured cells, primary tissues, and plasma. Previous publications have reported that up to ~90% of immunopurified peptides were specific for the MHC molecules. However, the yield of peptides from immunoprecipitation is relatively low (0.5–3%) due to the breakdown of complex and loss of peptides during the purification process [4]. We compared MAE and IP methods and found that they offered high overlap in identifications and some complementarity in the repertoire of peptides identified (Fig. 1c). Therefore, it is advantageous to combine both

methods, increasing the peptide sequence coverage of the immunopeptidome.

The additional challenge for MS-based immunopeptidomics technology is the generation of fragment ions for determining MHC peptide sequences. Compared to tryptic peptides, MHC peptides have much lower MS/MS identification rates due to their lack of predictable cleavage sites, shorter sequences, and lack of basic amino acid at C-termini. The lower ionization efficiency due to the absence of positive charges also negatively impacts the fragmentation of peptides and subsequent peptide identification by database search. Typically, the identification rate of MHC I peptides is about 10%, while an average of 50–60% identification rate is reported for tryptic peptides, either from collision-induced dissociation (CID) or high-energy collision dissociation (HCD) with hybrid LTQ-Orbitrap [5]. In recently published results, at least 30% of the total MHC I immunopeptides were identified based on singly charged ions [6]. Previous data also indicate that less than 30% of MS-identified MHC I peptides do not contain arginine (R) or lysine (K) residues [7, 8]. To address the challenges of peptide ionization and fragmentation, chemical derivatization of peptides can be employed. One derivatization strategy is to dimethylate amine-containing amino acids and N-termini, followed by alkylamidating carboxyl groups of acidic amino acids and C-termini [9]. Although dimethylation does not significantly alter the peptide fragmentation patterns, it increases the hydrophobicity of peptides such that it improves their chromatographic separation (Fig. 2a–c). Alkylamidation yields more y-ions in MS/MS fragmentation, increasing sequence coverage and confidence in identifying MHC peptides by database search (Fig. 2d–f). In this chapter, we describe the strategy of combining two isolation methods with chemical derivatization to improve MHC peptide identification. This approach provides an efficient way to expand the diversity of identified MHC I peptides as demonstrated by increased coverage of immunopeptidome from HCT 116 cells (Fig. 3).

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents. Prepare and store all reagents at room temperature (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing of waste materials.

2.1 Cell Culture

1. Cells: Human colon carcinoma cell line HCT 116 obtained from the American Type Culture Collection (ATCC).
2. Culture medium: RPMI 1640 medium with 10% fetal bovine serum and penicillin–streptomycin solution (final concentration of 1000 U/mL and 1 mg/mL).

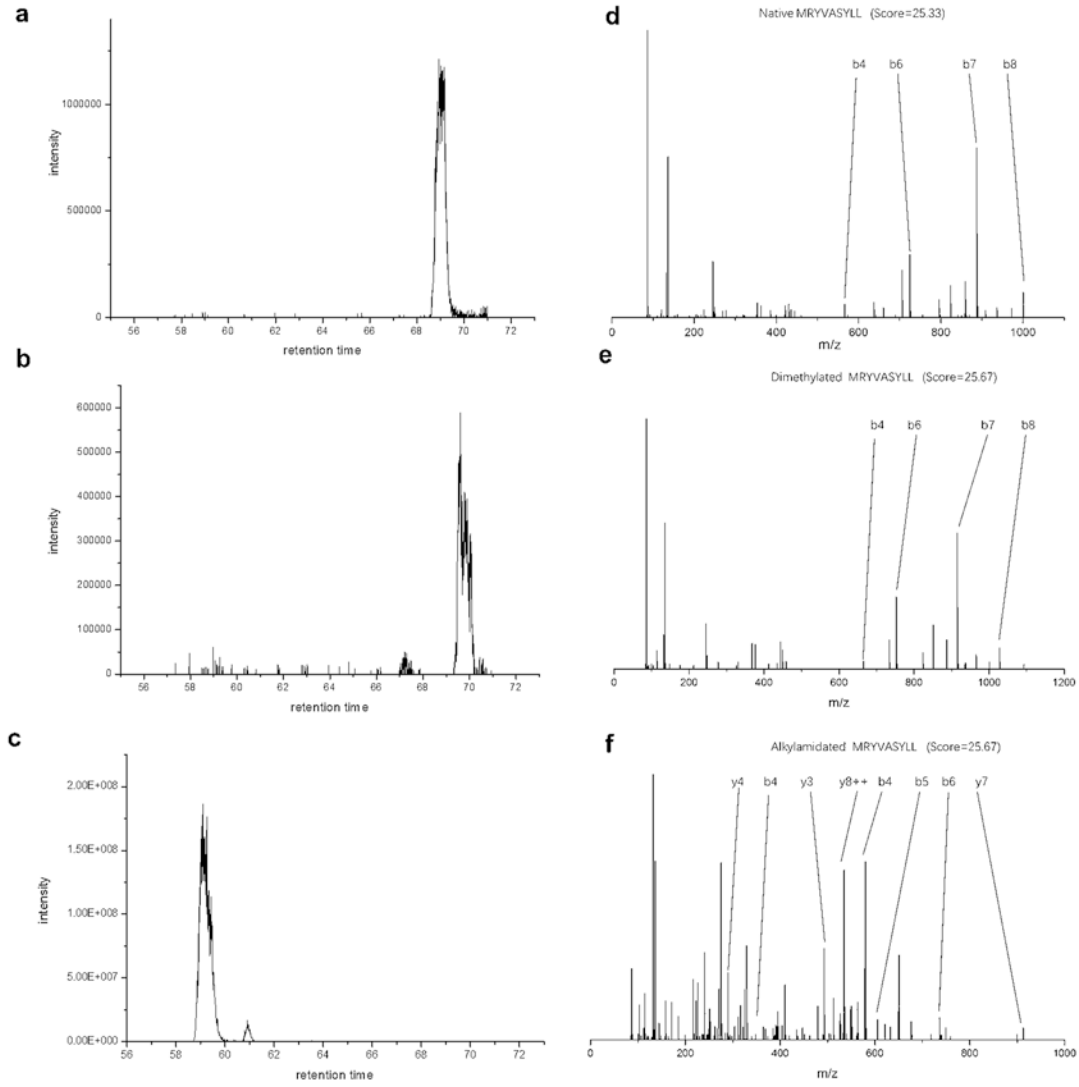


Fig. 2 LC-MS/MS identification of MHC peptides with chemical derivatization. (**a**) Chromatogram of native MRYVASYLL; (**b**) chromatogram of dimethylated MRYVASYLL; (**c**) chromatogram of alkylamidated MRYVASYLL; (**d**) MS/MS spectrum of native MRYVASYLL; (**e**) MS/MS spectrum of dimethylated MRYVASYLL; (**f**) MS/MS spectrum of alkylamidated MRYVASYLL

2.2 Immuno-precipitation

1. Cell lysis buffer: 200 mM sodium chloride, 20 mM Tris–HCl (pH 7.4), 0.5% CHAPS, 5 mM sodium vanadate and 5 mM sodium fluoride as phosphatase inhibitor, and 1× protease inhibitor cocktail (added freshly before each cell lysis).
2. W6/32 antibody: antibody purified from the conditioned media of HB-95 mouse hybridoma that recognizes human MHC I (HLA-A, HLA-B, HLA-C).

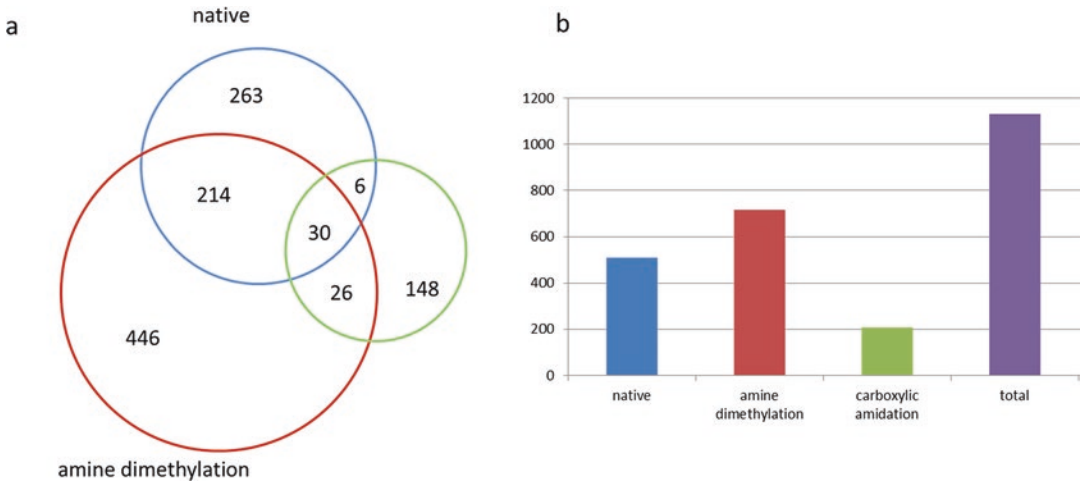


Fig. 3 Expanded coverage of MHC I immunopeptidome by chemical derivatization. (**a**) Overlap of identified unique peptides in different forms (native, dimethylated, and alkylamidated); (**b**) numbers of unique peptides identified in different forms

3. Cyanogen bromide (CNBr)-activated agarose or sepharose beads.
4. Phosphate-buffered saline (PBS).
5. Washing buffer A: 200 mM sodium chloride and 20 mM Tris–HCl (pH 7.4).
6. Blocking buffer: 100 mM Tris–HCl (pH 7.4).
7. Eluting solution: 10% acetic acid.
8. Antihuman HLA antibody.

2.3 Mild Acid Elution

1. Pre-made 1× phosphate-buffered saline (PBS).
2. Mild acid elution buffer: 0.131 M citric acid, 0.066 M Na_2HPO_4, NaCl 150 mM, pH 3.3.

2.4 MHC Peptide Purification

For MHC peptide purification, LC-MS grade water (e.g., Optima water from Fisher Scientific) is used to prepare all solutions.

1. Acidifying solution: 10% trifluoroacetic acid (TFA).
2. Elution buffer: 30% acetonitrile with 0.1% TFA.
3. Washing buffer B: 0.1% TFA.
4. C18 solid-phase extraction (SPE) cartridge.

2.5 Dimethylation of MHC Peptides

For MHC peptides dimethylation, LC-MS grade water is needed to prepare all the solutions.

1. Formaldehyde solution: 4% formaldehyde.
2. Reducing reagent: 100 mM sodium cyanoborohydride.

3. Dimethylation buffer: 100 mM triethylammonium acetate, pH 7.
4. Quenching solution A: 1% ammonia.
5. Acidifying solution: 10% TFA.
6. C18 SPE cartridge.

2.6 Alkylamidation of Dimethylated MHC Peptides

1. Morpholine.
2. Dimethyl sulfoxide.
3. (7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyAOP).
4. 3-(Dimethylamino)-1-propylamine (DMAPA).
5. Quenching solution B: 1% formic acid.
6. Chloroform.
7. C18 SPE cartridge.

2.7 LC-MS/MS Analysis

QExactive mass spectrometer coupled with nano-liquid chromatography (LC) system is recommended for the analysis of MHC peptides. The nano-LC column can be packed in-house with 1.9 μm reverse-phase beads (75 μm inner diameter (i.d.), length ≥ 20 cm) or columns that are commercially available, e.g., Acquity BEH C18 column from Waters, Acclaim PepMap 100 C18 LC columns from Thermo, etc.

2.8 Database Construction for Peptide Identification

BAM files from exome sequencing and RNA-seq data, GDC reference genome (GRCh38.d1.vd1.fa).

2.9 Database Search

1. Standard database: Uniprot human protein database (latest version).
2. Customized database: in-house-built protein database constructed using exome sequencing data.
3. Database search engine: Mascot (Matrix Science, version 2.3 or above).

3 Methods

3.1 Cell Culture

Grow HCT 116 cells in cell culturing medium at 37 °C with 5% CO_2 in T175 flask until 80% confluent. Ten flasks (175 cm^2) are sufficient for one MHC peptide isolation experiment (*see* **Note 1**).

3.2 Immunoprecipitation

1. Couple W6/32 antibody to CNBr-activated agarose or sepharose by incubation of antibody and beads overnight at 4 °C with rotation. Rinse the beads with PBS, and block the unreacted CNBr group with blocking buffer for three times.

The prepared beads can be stored in PBS at 4 °C (*see* **Notes 2** and **3**).

2. Wash the harvested cells with 10 mL of cold PBS for three times. Centrifuge at 600 × *g* for 3 min after each wash and discard supernatant.
3. Add 2 mL of cell lysis buffer with fresh protease inhibitor and phosphatase inhibitor to each flask, and detach the cells by gently shaking the flask.
4. Transfer the detached cells in lysis buffer to a Dounce homogenizer. Break the cells with loose piston and transfer the homogenate to a 50 mL tube (*see* **Note 4**).
5. Centrifuge the homogenate at 20,000 × *g* for 15 min and keep the supernatant.
6. Incubate the supernatant with 1 mL of agarose beads (W6/32 antibody conjugated) overnight at 4 °C with gentle rotation.
7. Centrifuge at 600 × *g* for 3 min to pellet the beads. Discard the supernatant, and wash the beads with the volume of washing buffer A, repeating three times.
8. Elute the bound MHC complex by adding 5 mL of eluting solution to the agarose beads and incubate at room temperature for 15 min. Centrifuge at 600 × *g* for 3 min and collect the supernatant (containing MHC peptides). This solution can be stored at −80 °C until LC-MS analysis.

3.3 Mild Acid Elution [3]

1. Wash the cells from Subheading 3.1 with 10 mL of PBS three times.
2. Add 5 mL of mild acid elution buffer to the flask, and gently rotate the flask to let the buffer cover all the cells for 1 min (*see* **Note 5**).
3. Transfer the buffer to clean tube, and centrifuge at 1000 × *g* to pellet detached cells and then at 15,000 × *g* for 15 min to remove debris. Store the elution at −80 °C until LC-MS analysis.

3.4 MHC Peptide Enrichment and Desalting

For MHC samples from both immunoprecipitation and mild acid elution, thaw the frozen samples at room temperature for enrichment, or use freshly prepared samples (*see* **Notes 6** and **7**).

1. Add acidifying solution to sample solution to adjust the final TFA concentration to 0.1% (i.e., 50 μL of acidifying solution to 5 mL of elution).
2. Activate the C18 cartridge with 1 mL of acetonitrile.
3. Clean the C18 cartridge with 1 ml of washing buffer B.
4. Load acidified sample solution to the C18 cartridge.

5. Wash the C18 cartridge with washing buffer B three times.
6. Elute the bound peptides with elution buffer.
7. Dry the elution down in a vacuum centrifuge.
8. Dissolve the dried peptides in washing buffer B and repeat the step from **2** to **7**. The MHC peptides are now ready for direct LC-MS/MS analysis or derivatization.

3.5 Dimethylation of Desalted MHC Peptides

Dimethylation reaction needs to be carried out in a fume hood due to the formation of toxic gas.

1. Dissolve desalted MHC peptides in 100 μL of the dimethylation buffer.
2. Add 5 μL of formaldehyde solution and 5 μL of reducing solution to MHC peptides, and vortex at room temperature for 1 h.
3. Quench the reaction by adding 20 μL of quenching solution A to the sample solution, and vortex at room temperature for 10 min.
4. Adjust the pH of sample solution to 2–3 with acidifying solution; leave sample in the hood for 30 min.
5. Desalt the dimethylated MHC peptides with procedure described in Subheading 3.4 (*see* **Note 8**).

3.6 Alkylamidation of Dimethylated MHC Peptides [9]

1. Dissolve dried dimethylated peptides in 5 μL of H_2O and 67 μL DMSO.
2. Add 3 mg of PyAOP, 7.6 μL DMAPA, and 2.6 μL of morpholine to the sample solution. Incubate for 4 h at room temperature with constant shaking.
3. Quench the reaction by adding 1 mL of quenching solution B.
4. Add 500 μL of chloroform to the sample solution and vortex for 1 min.
5. Centrifuge at 10,000 × *g* for 3 min and transfer the upper phase to a clean Eppendorf tube.
6. Add another 500 μL of chloroform to the sample solution and vortex for 1 min.
7. Centrifuge at 10,000 × *g* for 3 min and transfer the upper phase to a clean Eppendorf tube.
8. Dry the aqueous phase with a vacuum centrifuge.
9. Suspend the dried product in 1 mL washing buffer B and desalt the alkylamidated MHC peptides with the procedure described in Subheading 3.4.
10. Store the alkylamidated peptides at −80 °C until LC-MS analysis (*see* **Note 9**).

3.7 LC-MS/MS Analyses

1. Nano-LC separation: Solubilize dried peptides in 20 μL of 0.1% formic acid, and inject all of them to the trapping column. A 2-h gradient was performed from 5 to 35% acetonitrile (v/v) containing 0.1% formic acid at a flow rate between 200 and 500 nL/min (*see* **Note 10**).
2. Data-dependent MS/MS acquisition: The data acquisition method consisted of one full MS scan from m/z 300 to 1500, followed by data-dependent MS/MS scan for up to 12 most intense ions with a dynamic exclusion (repeat count of 2 and repeat exclusion duration of 30 s). Charge selection was enabled with selection of +2, +3 charged precursor for MS/MS. All data were recorded with the Xcalibur software and exported as .raw files for further analysis (*see* **Note 11**).

3.8 Database Building

1. Convert BAM files from exome sequencing and RNA-seq to FASTQ using biobambam [10] (*see* **Note 12**).
2. Align the DNA sequences against GDC reference genome (GRCh38.d1.vd1.fa) using BWA-MEM [11]. A manual for the alignment could be found at http://bio-bwa.sourceforge.net/bwa.shtml. RNA-seq data was aligned with STAR 2.5 [12] using the same genome and Gencode 22 annotations, following the manual (http://www.trii.org/courses/rnaseq_course_materials/STARmanual.pdf).
3. Variant calls were obtained using VarDictJava [13]. VarDict simultaneously calls single-nucleotide variant (SNV), multi-nucleotide variant (MNV), small insertions and deletions (InDels), and complex and structural variants, expanding the detected genetic driver landscape of tumors. The manual for the Java re-implementation can be found at https://github.com/AstraZeneca-NGS/VarDictJava. The effect of variants was evaluated using the Variant Effect Predictor [14] (https://useast.ensembl.org/info/docs/tools/vep/index.html). Variants with moderate or high impact, inside exons of protein-coding genes, were retained.
4. Mutated protein sequences were generated with one sequence per variant through online translation tools, e.g., Expasy (https://web.expasy.org/translate/), blastx from NCBI or EBI (https://www.ebi.ac.uk/Tools/st), etc.

3.9 Database Search and Data Processing

1. Data files acquired from MS were transformed to .mgf files using ProteoWizard [15] (*see* **Note 13**).
2. Transformed files were searched using Mascot (Matrix Science Inc.) against the most up-to-date Uniprot human database with reversed sequence as decoy database. Mass tolerance was set as 8 ppm for precursor and 0.05 Da for fragment ions (*see* **Note 14**).

3. For native peptides, only methionine oxidation (+15.994915) was selected as variable modification.
4. For dimethylated peptides, oxidation and dimethylation at lysine and peptide N-terminus (+28.031300) were set as variable modification.
5. For peptides with derivatization of carboxylic acid, dimethylation at lysine and peptide N-terminus (+28.031300) was set as fixed modification. Methionine oxidation and alkylamidation of aspartic acid, glutamic acid, and C-terminal peptides (+84.105130) were set as variable modification.
6. After database searching, peptide-spectrum matches (PSM) were filtered with the following criteria: top 1 match, Mascot score ≥ 22, and 8 to 14 amino acids in length.
7. False discovery rate (FDR) was calculated as follows: $FDR = [(PSM)_{decoy} \times 2]/[(PSM)_{target} + (PSM)_{decoy}]$. An FDR value between 5 and 9% would be expected (*see* **Note 15**).

4 Notes

1. The amount of cells used for each MHC peptide isolation can be increased up to 1×10^9 (usually $2–5 \times 10^8$ are used for each experiment). It is recommended to start with a lower amount of starting material to optimize the ratio of antibody to cells and the washing/elution conditions.
2. W6/32 antibody can also be conjugated to aldehyde-activated agarose/sepharose beads through reductive amidation. The chemical bond from aldehyde-activated agarose/sepharose is more stable than that from CNBr-activated agarose/sepharose and has no net charge. This results in less nonspecific binding in affinity purification procedures. Immunoprecipitation can also use protein A agarose beads for immobilizing W6/32 antibody by cross-linking. The usage of immobilized affinity column avoids the co-elution of the antibody, which may block the LC column during LC-MS/MS analysis.
3. To evaluate the conjugation efficiency before immunoprecipitation, collect the supernatant after coupling reaction and compare the protein concentration with the original antibody solution. Low conjugation efficiency may lead to poor immunoprecipitation. When using newly prepared materials, test with a small amount of cell lysate, and use Western blotting to check the efficiency of immunoprecipitation. Before LC-MS/MS analysis, always use Western blotting as quality control.
4. Avoid using hash detergents like sodium dodecyl sulfate (SDS) for cell lysis as the strong interaction between SDS and protein may disrupt the MHC complex, resulting in no affinity with

W6/32 antibody. For similar reasons, we do not recommend the use of sonication for cell lysis.

5. During mild acid elution, adjust the speed of shaking of the cell culture flask to avoid cells detaching. The speed is dependent on the adherence of different cell lines. If there are too many detached cells in the mild acid elution buffer, discard the samples, and repeat the mild acid elution with a fresh flask of cells.
6. Contamination from HLA and beta-2 microglobulin is an issue for either immunoprecipitation or mild acid elution. The intact proteins may cause LC column blockage during LC-MS/MS analysis. An alternative is to use Amicon molecular weight cutoff filtering membrane to remove those protein contaminants. Although it is efficient in filtering proteins, the process can also bind some peptides through nonspecific interaction.
7. Both immunoprecipitation and mild acid elution are reproducible for isolating MHC peptides from cells. As demonstrated in Fig. 1, these two methods are complementary to each other, and the combination of both methods would increase the coverage of immunopeptidome and discover novel neoantigens. More importantly, using two methods together would increase the confidence in identifying MHC peptides with posttranslational modifications. However, mild acid elution is not applicable to tissue samples.
8. Peptides need to be completely desalted as residual Tris can interfere with the dimethylation reaction. Incomplete dimethylation would cause failure in the following alkylamidation step. Run an LC-MS/MS analysis and database search to check the efficiency of dimethylation before proceeding to the alkylamidation. A minimum of 95% of identified peptides should be dimethylated.
9. The amount of reagent, especially DMAPA, is critical for the completion of alkylamidation. The optimal amount of DMAPA can be evaluated with the peptides purified from the same number of cells. Other tertiary amine reagents could also be used for alkylamidation.
10. The suitability of the LC-MS/MS system can be examined by running a tryptic digestion of bovine serum albumin (1 picomole), and a sequence coverage of 85% is needed.
11. The data-dependent setting with a Q-TOF instrument might be different from the parameters described above. Taking Synapt G2 from Waters, for example, only three most abundant precursors are selected for MS/MS, and the. RAW files need to be processed and transformed to .mgf by PLGS.

12. The sequencing data of NCI 60 human cancer cell lines can be downloaded from CellMiner (https://discover.nci.nih.gov/cellminer/). For immortalized B cells or tissue samples from patients, exome sequencing or RNA-seq needs to be conducted, and the HLA genotype needs to be mapped.
13. ProteoWizard transforms raw data from mass spectrometry to formats that are recognized by database search engines (e.g., .mgf, mzxml, etc.). .mgf files are suitable for Mascot search. For the transformation, MS level from 1 to 2 needs to be selected. The selection of activation method is dependent on the instrument used, e.g., high-energy collision dissociation (HCD) for QExactive and collision-induced dissociation (CID) for Q-TOF like Synapt G2.
14. Other database search engines (i.e., Maxquant, X!Tandem, MSGF+, etc.) can also be used. Combining the results from different engines could increase the coverage and confidence in identifying neoantigens.
15. Distinct from mainstream proteomics, a loose false discovery rate (FDR) criterion can be set to increase the sensitivity in identifying MHC peptides. MHC peptides have short sequences, thus offering fewer matched fragments from database search and resulting in lower scores. Setting nonenzymatic digestion also causes lower scores compared with tryptic peptides. The false positives can be filtered off by the length of amino acids (8–14 for MHC I peptide) and the predicted binding affinity to particular HLA genes. MHC peptides identified with mutations are likely to be neoantigens if passing the filtering criteria (e.g., length and affinity). It is strongly recommended to validate the peptide sequences before any immunogenicity assays. The existence of a targeted peptide can be validated by LC-MS/MS analysis of synthetic peptide with the same sequence to examine its retention time in reversed-phase chromatography and expected MS/MS spectra.

References

1. Schumacher TN, Schreiber RD (2015) Neoantigens in cancer immunotherapy. Science 348:69–74
2. Gubin MM, Artyomov MN, Mardis ER, Schreiber RD (2015) Tumor neoantigens: building a framework for personalized cancer immunotherapy. J Clin Invest 125:3413–3421
3. Fortier MH, Caron E, Hardy MP, Voisin G, Lemieux S, Perreault C et al (2008) The MHC class I peptide repertoire is molded by the transcriptome. J Exp Med 205:595–610
4. Hassan C, Kester MG, Oudgenoeg G, de Ru AH, Janssen GM, Drijfhout JW et al (2014) Accurate quantitation of MHC-bound peptides by application of isotopically labeled peptide MHC complexes. J Proteome 109:240–244
5. Mommen GPM, Frese CK, Meiring HD, van Gaans-van den Brink JV, de Jong A, van Els C et al (2014) Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). Proc Natl Acad Sci U S A 111:4507–4512

6. Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ, Mann M (2015) Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. Mol Cell Proteomics 14:658–673
7. Bassani-Sternberg M, Braunlein E, Klar R, Engleitner T, Sinitcyn P, Audehm S et al (2016) Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. Nat Commun 7:13404
8. Laumont CM, Daouda T, Laverdure JP, Bonneil E, Caron-Lizotte O, Hardy MP et al (2016) Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. Nat Commun 7:10238
9. Frey BL, Ladror DT, Sondalle SB, Krusemark CJ, Jue AL, Coon JJ et al (2013) Chemical derivatization of peptide carboxyl groups for highly efficient electron transfer dissociation. J Am Soc Mass Spectrom 24:1710–1721
10. Tischler G, Leonard S (2014) Biobambam: tools for read pair collation based algorithms on BAM files. Source Code Biol Med 9:13
11. Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997
12. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S et al (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21
13. Lai ZW, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R et al (2016) VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 44:e108
14. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A et al (2016) The Ensembl variant effect predictor. Genome Biol 17:122
15. Kessner D, Chambers M, Burke R, Agusand D, Mallick P (2008) ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–2536



Chapter 16

Enrichment of Phosphorylated MHC Peptides with Immobilized Metal Affinity Chromatography and Titanium Dioxide Particles

Rui Chen and Jianjun Li

Abstract

Phosphorylation is one of the most important forms of posttranslational modification. Dysregulation of phosphorylation is implicated in tumorigenesis, with cancerous signaling pathways activated by kinases. For immunotherapy with neoantigen-based peptides, phosphopeptides derived from aberrantly phosphorylated proteins presented by major histocompatibility complex (MHC) are promising candidates due to their specificity to elicit cytotoxic T-cell responses. Unlike other MHC peptides, phosphorylated MHC peptides cannot be predicted from DNA sequences, and their identification relies on the direct detection of phosphopeptides using mass spectrometry (MS). For MS detection, it is extremely important to enrich phosphorylated peptides from the complex repertoire of MHC peptides. Herein, we describe the combined use of immobilized metal affinity chromatography and titanium dioxide nanoparticles for phosphopeptides enrichment from immunopeptidome.

Key words Immunotherapy, MHC peptides, Phosphorylation, IMAC, LC-MS/MS

1 Introduction

Protein phosphorylation by kinases and dephosphorylation by phosphatases are key components of protein signaling in cellular networks [1]. It was estimated that phosphorylation affects around 30% of the proteome and regulates many cellular processes [2]. Dysregulation of protein phosphorylation has been found to be a unique feature of cancer [3], with different phosphorylation pathways observed to be activated in tumors. Quantitative analysis of altered phosphorylation will provide unique insights into cancer, the identification of cancer-specific biomarkers and discovery of new targets of cancer therapy [4–6]. With the latest advances in mass spectrometry (MS) techniques, especially enhanced scanning speed and detection sensitivity, phosphoproteomics has revealed tumor-associated signatures from various samples, such as cancer cell lines [7], cancer stem cells [8], patient-derived xenograft [9],

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_16, © Springer Science+Business Media, LLC, part of Springer Nature 2019

and tumor tissues [10]. Phosphorylation is more effective in triggering an immune response from cytotoxic T cells than unmodified peptides [6, 11]. Identification of phosphorylated peptides bound to the major histocompatibility complex (MHC) is important in the development of cancer immunotherapy with neoantigens. Phosphorylated MHC peptides have been detected directly from LC-MS/MS analysis of whole immunopeptidome isolated from cells or tissues with immunoprecipitation [12, 13]. However, the low abundance of phosphorylated MHC peptides in comparison to their non-modified counterparts makes their detection and identification challenging. Thus, enrichment of phosphopeptides prior to LC-MS/MS analysis is crucial to increase the specificity and accuracy in identification of phosphorylated MHC peptides.

There are several published protocols for phosphopeptide enrichment that are compatible with downstream MS analyses. The majority of them are based on immobilized metal affinity chromatography (IMAC) or metal oxide affinity chromatography (MOAC) [14]. Iron chelated to nitrilotriacetic was the first IMAC exploited for phosphopeptide enrichment. This was superseded by metal oxide particles and Ti^{4+} IMAC [15–17]. TiO_2 and Ti^{4+} IMAC enrichment methods have a higher affinity for phosphopeptides than that of Fe^{3+}-based IMAC. Unfortunately, the Ti^{4+} IMAC used in their report is not available to the public. Most large-scale studies still use TiO_2 for the enrichment due to its high reproducibility. Recently, improvement in Fe^{3+} IMAC suggested that nearly 100% specificity for phosphopeptides could be achieved [18]. The literature suggests that the coverage of the phosphoproteome was primarily determined by the metal ion used for enrichment. Some reports show that different methods provided varied immunopeptidome coverage [19].

In this protocol, we present a complementary enrichment of phosphorylated MHC peptides with Fe^{3+} IMAC and TiO_2 nanoparticles. This strategy provides a mechanism to enhance phosphopeptide coverage without sacrificing sensitivity and specificity. The proposed method was evaluated by analyzing immunopeptidome from A549 cell lines, and complementary identification of phosphorylated MHC peptides was achieved (*see* Table 1).

2 Materials

2.1 Cell Culture

1. Cells: Human lung cancer cell line A549 obtained from the American Type Culture Collection (ATCC).
2. Culture media: Dulbecco's Modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) and penicillin-streptomycin solution (final concentration of 1000 U/mL and 1 mg/mL, respectively).

Table 1
Example of phosphorylated MHC peptides identified with IMAC and TiO_2 from A549 cells

Sequence	Protein ID	IMAC	TiO_2	Mascot score
SVTEHPS#PSPLLS	P07910	√		34.1
Y#TEEEEEE	Q6P5Q4	√		38
PDS#DWEPEE	Q9NPE2	√		30.3
FHS#EENGS	Q9Y4A5		√	23.6
S#GGGDSERAAALS#P	Q8WZA9	√		25.7
CDPNS#ICT#NTVGSFK	P49746	√		24
FQIS#QDEDESESED	Q6P5R6	√	√	47.3
DS#EEDEEHTIITDT	P10646		√	33.3

2.2 Immuno-precipitation

1. Cell lysis buffer: 200 mM sodium chloride, 20 mM Tris–HCl (pH 7.4), 0.5% CHAPS, 5 mM sodium vanadate, and 5 mM sodium fluoride as a phosphatase inhibitor, 1× protease inhibitor cocktail (add before each cell lysis).
2. W6/32 antibody: antibody purified from the conditioned media of HB-95 mouse hybridoma that recognizes human MHC I (HLA-A, B, C).
3. Cyanogen bromide (CNBr)-activated agarose or sepharose beads.
4. Phosphate buffered saline (PBS): purchased pre-prepared.
5. 100 mM Tris(hydroxymethyl)aminomethane hydrochloride (100 mM Tris–HCl).
6. 10% (v/v) acetic acid.

2.3 MHC Peptide Preconcentration and Desalting

For MHC peptide purification, LC-MS grade water is needed to prepare all solutions:

1. Acidifying solution: 10% (v/v) trifluoroacetic acid (TFA).
2. Elution buffer: 30% (v/v) acetonitrile with 0.1% (v/v) TFA.
3. Washing buffer: 0.1% (v/v) TFA.
4. C18 solid phase extraction (SPE) cartridge.

2.4 Enrichment of Phosphorylated MHC Peptides

All reagents must be analytical or higher grade unless specified.

1. MHC immunopeptidome isolated from A549 cells.
2. Nitrilotriacetic acid (NTA) silica with a particle size of 16–24 μm.
3. Iron (III) chloride ($FeCl_3$).

4. Titanium dioxide nanoparticles, 20 nm in particle size.
5. C18 spin column.
6. Acetonitrile (ACN), LC-MS grade.
7. Water, LC-MS grade.
8. Dihydroxybenzoic acid (DHB).
9. Formic acid (FA), LC-MS grade.
10. Ethylenediaminetetraacetic acid (EDTA), analytical grade.
11. Trifluoroacetic acid (TFA), analytical grade.
12. 25% ammonia solution in water.
13. Acetic acid, analytical grade.
14. IMAC resin washing buffer 1: 0.01% (v/v/) acetic acid.
15. IMAC resin washing buffer 2: 0.01% (v/v) acetic acid, 15 ACN.
16. Fe^{3+} IMAC loading buffer: 0.1 M acetic acid.
17. Fe^{3+} IMAC washing buffer: 30% (v/v) ACN, 0.07% (v/v) TFA.
18. Fe^{3+} IMAC eluting buffer: 0.5% (v/v) ammonia.
19. TiO_2 loading buffer: 80% (v/v) ACN, 6% (v/v) TFA and 50 mg/mL DHB.
20. TiO_2 washing buffer: 50% (v/v) ACN, 0.1% (v/v) TFA.
21. TiO_2 eluting buffer: 5% (v/v) ammonia.

2.5 LC-MS/MS Analysis

Q Exactive mass spectrometer coupled with a nano-LC is recommended for the analysis of MHC peptides. The nano-LC column can be packed in-house with 1.9 μm reverse phase beads (75 μm i.d., length ≥ 20 cm) or columns that are commercially available, e.g., Acquity BEH C18 column from Waters, Acclaim PepMap 100 C18 LC columns from Thermo, etc.

2.6 Database Search

1. Standard Protein database: UniProt human protein database.
2. Customized database: In-house-built protein database constructed using exome sequencing data [20].
3. Database search engine: Mascot (Matrix Science, version 2.3 or later).

3 Methods

3.1 Cell Culture

Grow A549 cells in cell culture media at 37 °C with 5% CO_2 in T175 flask until 80% confluent. Collect cells from ten flasks for one MHC peptide isolation experiment (*see* **Note 1**).

3.2 Immunoprecipitation

1. Couple W6/32 antibody to CNBr-activated agarose or sepharose by incubation of antibody and beads overnight at 4 °C. Overnight incubation must be with rotation. Rinse the beads, and block the unreacted CNBr group with 100 mM Tris–HCl buffer three times. The prepared beads can be stored in PBS at 4 °C.
2. Add 10 mL of cold PBS to harvested cells. Centrifuge at 600 ×*g* for 3 min wash and discard supernatant. Repeat a total of three times.
3. Add 2 mL of cell lysis buffer with fresh protease inhibitor and phosphatase inhibitor to each flask, and detach the cells by gently shaking the flask.
4. Transfer the detached cells in lysis buffer to a glass homogenizer. Break the cells with loose piston, and transfer the homogenate to a 50 mL tube (*see* **Note 2**).
5. Centrifuge the homogenate at 20,000 ×*g* for 15 min and keep the supernatant. Store 100 μL of lysate for western blotting (*see* **Note 3**), and use the rest for immunoprecipitation.
6. Incubate the lysate with 1 mL of agarose beads (W6/32 antibody conjugated) overnight at 4 °C with gentle rotation.
7. Centrifuge at 600 ×*g* for 3 min to pellet the beads. Discard the supernatant and wash the beads with the same volume of washing buffer as that used for cell lysis. Repeat wash three times.
8. Elute the bound MHC by adding 5 mL of 10% acetic acid to the agarose beads, and incubate at room temperature for 15 min. Centrifuge at 600 ×*g* for 3 min and collect the supernatant as elution. If desired, retain 100 μL of the elution for western blotting confirmation of MHC isolation, and store the rest in −80 °C until LC-MS analysis (*see* **Note 3**).

3.3 MHC Peptide Preconcentration and Desalting

1. Add acidifying solution to sample solution to adjust TFA concentration at 0.1% (i.e., 50 μL of acidifying solution to 5 mL of elution).
2. Activate the C18 cartridge with 1 mL of ACN.
3. Clean the C18 cartridge with 1 mL of washing buffer.
4. Load acidified sample solution from **step 1** to the C18 cartridge.
5. Wash the C18 cartridge with 1 mL washing buffer three times.
6. Elute the bound peptides with 1 mL of elution buffer.
7. Evaporate eluted peptides to dryness using a Speed Vac.
8. Dissolve the dried peptides in 1 mL washing buffer and repeat the steps from **2** to **7**. The MHC peptides are ready for phosphopeptide enrichment.

3.4 Preparation of Iron (III)-NTA IMAC

1. Reconstitute the NTA silica with water (1 mg/mL) in a 1.5 mL Eppendorf tube.
2. Centrifuge at 13,000 × *g* for 5 min, and discard the supernatant.
3. Add 1 mL of 50 mM ETDA and vortex for 5 min at room temperature.
4. Centrifuge at 13,000 × *g* for 5 min, discard the supernatant.
5. Repeat **step 2** twice.
6. Activate the resin by adding 1 mL of 100 mM $FeCl_3$ and stirring overnight.
7. Centrifuge at 13,000 × *g* for 5 min, discard the supernatant.
8. Wash the resin with 1 mL IMAC resin washing buffer 1, followed by wash with IMAC resin washing buffer 2 and finally with another wash with IMAC resin washing buffer 1.
9. Store the Fe-NTA IMAC resin in 0.01% acetic acid at 4 °C at the concentration of 50 mg/mL. The resin can be stored for up to 1 month.

3.5 Enrichment of Phosphorylated MHC Peptides with IMAC

1. Reconstitute immune-purified HLA-associated peptide samples in 100 μL Fe^{3+} IMAC loading buffer.
2. Add 20 μL of Fe^{3+} IMAC suspension to the MHC peptide sample solution, and vortex for 30 min (*see* **Note 4**).
3. Centrifuge at 13,000 × *g* for 5 min, discard the supernatant.
4. Add 500 μL of Fe^{3+} IMAC washing buffer and vortex for 15 min (*see* **Note 5**).
5. Centrifuge at 13,000 × *g* for 5 min, discard the supernatant.
6. Adding 100 μL of Fe^{3+} IMAC elution buffer and vortex for 15 min.
7. Centrifuge at 13,000 × *g* for 5 min, transfer the supernatant to a clean tube.
8. Add 100 μL of acidifying solution.
9. Activate C18 spin column by adding 50 μL of ACN, spin at 1000 × *g* for 1 min.
10. Wash the spin column with 50 μL of washing buffer.
11. Load the peptides onto the spin column, spin at 1000 × *g* for 1 min.
12. Wash the spin column with 50 μL of washing buffer.
13. Elute desalted peptides with 50 μL elution buffer.
14. Dry the peptides eluted from the C18 column using a Speed Vac. Store the dried samples at −80 °C until LC-MS analysis.

3.6 Enrichment of Phosphorylated MHC Peptides with TiO_2 Nanoparticles

1. Preparation of TiO_2 nanoparticles suspension of 20 mg/mL in 6% TFA, 80% acetonitrile (*see* **Notes 6** and **7**).
2. Reconstitute purified MHC peptides in 200 μL of loading buffer.
3. Adding 10 μL of TiO_2 suspension to the MHC peptides solution.
4. Shake at room temperature for 30 min.
5. Centrifuge at 10,000 × *g* for 5 min, discard supernatant.
6. Add 500 μL of washing buffer and vortex for 15 min.
7. Centrifuge at 10,000 × *g* for 5 min, discard supernatant.
8. Add 200 μL of 5% ammonia, shake for 15 min.
9. Centrifuge at 10,000 × *g* for 5 min, transfer to supernatant to a clean tube without disturbing the pellet.
10. Acidify eluted peptides solution with 100 μL of acidifying solution.
11. Desalt the eluted peptides with C18 spin column as described in Subheading 3.5.
12. Dry the eluted peptides using a Speed Vac, and store the samples at −80 °C until MS analysis.

3.7 LC-MS/MS Analyses

1. Nano-LC separation: A 2-h gradient was performed from 5% to 35% acetonitrile (v/v) containing 0.1% formic acid (v/v) at a flow rate between 200 nL/min and 500 nL/min.
2. Data-dependent MS/MS acquisition: The data acquisition method consisted of one full MS scan from m/z 300 to 1500, followed by data-dependent MS/MS scan for up to 12 most intense ions with a dynamic exclusion (repeat count of two and repeat exclusion duration of 30 s). Charge selection was enabled to select +2 and +3 charged precursors for MS/MS. All data were recorded with the Xcalibur software and exported as .raw files for further analysis.

3.8 Database Search and Data Processing

1. Raw files were transformed to .mgf files with ProteoWizard [21].
2. Search with Mascot (Matrix Science Inc.) against the latest UniProt human database with reversed sequence as decoy database. Mass tolerance was set as 8 ppm for precursor and 0.05 Da for fragment ions. No enzyme digest was selected.
3. Set methionine oxidation (+15.994915) and phosphorylation at serine, threonine, and tyrosine (+79.966331) as variable modification.

4. After database searching, peptides-spectra matches (PSM) were filtered with the following criteria: top 1 match, Mascot Score ≥ 22, and 8–14 amino acids in length. False discovery rate (FDR) was calculated as the following: FDR = $[(PSM)_{decoy} \times 2]/[(PSM)_{target} + (PSM)_{decoy}]$. An FDR value between 5 and 9% would be expected (*see* **Notes 8** and **9**).

4 Notes

1. The amount of cells used for phosphopeptide enrichment can be increased up to 1×10^9, to ensure sufficient starting material for MS detection of phosphopeptides. For the optimization of the phosphopeptide enrichment procedure, the starting amount could be lower. However, when working with tissue samples, the amount of starting material is sample-dependent as the expression of HLA from each sample has large variation.
2. Phosphatase inhibitors are added to the lysis buffer for phosphopeptide enrichment to prevent dephosphorylation of peptides. Phosphatase inhibitor cocktail is also available from a number of commercial sources.
3. We suggest retaining 100 μL of the elution for SDS-PAGE and western blotting to confirm success of immuno pulldown of MHCs. We suggest running an SDS-PAGE of the original cell lysate, flow-through, wash, and elution from immunoprecipitation. After standard western blot transfer to membranes, we typically probe with 1:2000 diluted antihuman HLA antibody overnight at 4 °C with constant rotation and then 1:5000 diluted secondary antibody for detection. The intensity of blots from flow-through, wash, and elution is compared to verify the efficiency of immunoprecipitation.
4. Off-line SPE was used in this protocol. However, it has been reported that the specificity of enrichment could be increased by packing the material in a spin column or using a high-performance liquid chromatograph (HPLC). We chose off-line SPE as it is easy to operate and does not require specialized instrumentation.
5. For both IMAC and MOAC, the volume of washing buffer and number of washing steps might need to be optimized for each sample. This is dependent upon the complexity of the samples. Extra washing is needed if the specificity of phosphopeptide identification is low. Since immunopeptidome is less complex than tryptic proteome, only one wash is used in this protocol to minimize sample loss.

6. TiO_2 with larger particles sizes (up to 50 μm) can be used for phosphopeptide enrichment. Other metal oxide, e.g., ZrO_2, may also provide complementary coverage.
7. Lactic acid and glycolic acid are alternatives for DHB in the loading buffer to reduce the non-specific binding of peptides containing acidic amino acids.
8. The presence of phosphate group can be verified by the neutral loss in MS/MS spectrum, but the intensity is dependent on peptide sequence.
9. A lower FDR setting (e.g., 5%) can be used as initial PSM filter, and extra criteria (e.g., peptide length, binding affinity to HLA) need to remove possible false identifications. For lower PSM scores, we suggest manual inspection of the MS/MS spectra to confirm amino acid sequences.

References

1. Hunter T (2000) Signaling—2000 and beyond. Cell 100:113–127
2. Ubersax JA, Ferrell JE Jr (2007) Mechanisms of specificity in protein phosphorylation. Nat Rev Mol Cell Bio 8:530–541
3. Brognard J, Hunter T (2011) Protein kinase signaling networks in cancer. Curr Opin Genet Dev 21:4–11
4. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA et al (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 143:1174–1189
5. Francavilla C, Lupia M, Tsafou K, Villa A, Kowalczyk K, Jersie-Christensen RR et al (2017) Phosphoproteomics of primary cells reveals druggable kinase signatures in ovarian cancer. Cell Rep 18:3242–3256
6. Mohammed F, Cobbold M, Zarling AL, Salim M, Barrett-Wilt GA, Shabanowitz J et al (2008) Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. Nat Immunol 9:1236–1243
7. Locard-Paulet M, Lim L, Veluscek G, McMahon K, Sinclair J, Van Weverwijk A et al (2016) Phosphoproteomic analysis of interacting tumor and endothelial cells identifies regulatory mechanisms of transendothelial migration. Sci Signal 9:ra15-ra15
8. Yi T, Zhai B, Yu Y, Kiyotsugu Y, Raschle T, Etzkorn M et al (2014) Quantitative phosphoproteomic analysis reveals system-wide signaling pathways downstream of SDF-1/CXCR4 in breast cancer stem cells. Proc Natl Acad Sci U S A 111:E2182–E2190
9. K-L H, Li S, Mertins P, Cao S, Gunawardena HP, Ruggles KV et al (2017) Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. Nat Commun 8:14864
10. Dazert E, Colombi M, Boldanova T, Moes S, Adametz D, Quagliata L et al (2016) Quantitative proteomics and phosphoproteomics on serial tumor biopsies from a sorafenib-treated HCC patient. Proc Natl Acad Sci U S A 113:1381–1386
11. Zarling AL, Polefrone JM, Evans AM, Mikesh LM, Shabanowitz J, Lewis ST et al (2016) Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. Proc Natl Acad Sci U S A 103:14889–14894
12. Bassani-Sternberg M, Bräunlein E, Klar R, Engleitner T, Sinitcyn P, Audehm S et al (2016) Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. Nat Commun 7:13404
13. Laumont CM, Daouda T, Laverdure J-P, Bonneil É, Caron-Lizotte O, Hardy M-P et al (2016) Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. Nat Commun 7:10238
14. Olsen JV, Mann M (2013) Status of large-scale analysis of post-translational modifications by mass spectrometry. Mol Cell Proteomics 12:3444–3452

15. Andersson L, Porath J (1986) Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. Anal Biochem 154:250–254
16. Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ (2004) Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Anal Chem 76:3935–3943
17. Zhou H, Low TY, Hennrich ML, van der Toorn H, Schwend T, Zou H et al (2011) Enhancing the identification of phosphopeptides from putative basophilic kinase substrates using Ti (IV) based IMAC enrichment. Mol Cell Proteomics 10:M110.006452
18. Potel CM, Lin M-H, Heck AJ, Lemeer S (2018) Defeating major contaminants in Fe3+-IMAC phosphopeptide enrichment. Mol Cell Proteomics 17:1028–1034
19. Abelin JG, Trantham PD, Penny SA, Patterson AM, Ward ST, Hildebrand WH et al (2015) Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. Nat Protoc 10:1308–1318
20. Chen R, Fauteux F, Foote S, Stupak J, Tremblay T-L, Gurnani K et al (2018) Chemical derivatization strategy for extending the identification of MHC class I immunopeptides. Anal Chem 90:11409–11416
21. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S et al (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30:918–920



Chapter 17

Mass Spectrometric Identification and Molecular Modeling of Glycopeptides Presented by MHC Class I and II Processing Pathways

Stacy A. Malaker and Michael J. Ferracane

Abstract

Aberrant glycosylation is a hallmark of cancer that contributes to the disease's ability to evade the immune system. As the MHC processing pathways communicate cellular health to circulating CD8+ and CD4+ T-cells, MHC-associated glycopeptides are likely a source of neoantigens in cancer. In fact, recent advances in mass spectrometry have allowed for the detection and sequencing of tumor-specific glycopeptides from the MHC class I and class II processing pathways. Here, we describe methods for detecting, sequencing, and modeling these MHC-associated glycopeptides.

Key words Mass spectrometry, MHC-associated glycopeptides, Glycopeptide analysis, Neoantigens, Molecular modeling

1 Introduction

Glycosylation is a ubiquitous and diverse posttranslational modification (PTM). Three main types of glycosylation exist in humans: (1) extracellular N-linked glycosylation, (2) extracellular O-linked glycosylation, and (3) intracellular O-linked β-N-acetylglucosamine (O-GlcNAc). All of these glycosylation patterns are altered in cancer, either through hypersialylation, alteration of branching patterns, truncation of structures, or global upregulation [1].

The major histocompatibility complex (MHC) class I and II processing pathways communicate cellular health to the immune system, and several groups have hypothesized that glycopeptides could affect T-cell recognition in dysregulated cells. As such, numerous studies have attempted to understand if glycopeptides are presented on the cell surface via these pathways and how they affect activation or inhibition of T-cells. For instance, synthetic O-GlcNAcylated peptides have been shown to bind MHC class I complexes and elicit specific T-cell responses in mice [2–4]; X-ray

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_17, © Springer Science+Business Media, LLC, part of Springer Nature 2019

structures confirm that the O-GlcNAc moiety is solvent exposed and available for interaction with the CD8+ T-cell receptor [5]. Additionally, N-linked carbohydrates have been shown to be important for CD4+ T-cell reactivity in arthritis and melanoma [6–8]. Despite these efforts, very few glycopeptides naturally presented by either the MHC class I or II processing pathways have ever been reported [9, 10].

Recent advances in mass spectrometry have finally made it possible to detect and characterize these low-level analytes. One of the most challenging aspects of this method is obtaining quality fragmentation (MS2) spectra of trace glycopeptides present among more abundant, unmodified counterparts. But now, through parallelization of spectra acquisition, the Thermo Orbitrap Tribrid series of mass spectrometers is able to acquire more MS2 spectra than previously obtainable, extending the detection limit immensely. This analytical improvement, alongside advances in glycoproteomic search algorithms [11], makes detection and sequencing of MHC-associated glycopeptides relatively routine.

Here, we describe the methods that allowed us to detect and sequence hundreds of MHC-associated class I and II glycopeptides [12, 13]. We also outline methods used to model how some of these glycopeptides likely bind to MHC class II molecules and potentially evade detection by the immune system.

2 Materials

Prior to beginning the wet lab procedures, prepare all solutions in new glassware. All solutions are stored at room temperature unless otherwise noted.

2.1 Reagents

1. Glacial acetic acid, >99.9% purity.
2. High-performance liquid chromatography (HPLC)-grade acetonitrile, >99.8% purity.
3. Liquid chromatography-mass spectrometry (LC-MS) grade water.
4. Isopropanol, 99.9% purity.
5. Reverse-phase HPLC buffer A: 0.1 M acetic acid in water.
6. Reverse-phase HPLC buffer B: 0.1 M acetic acid in 70% acetonitrile, 30% water (v/v). Sonicate for 20 min before use.
7. Peptide standard solution: 500 fmol/μL angiotensin I human acetate hydrate (Sigma-Aldrich, >90% purity) and 500 fmol/μL vasoactive intestinal peptide fragment 1–12 (Sigma-Aldrich, >97% purity) in 0.1 M acetic acid. Store at −35 °C until needed.

2.2 Equipment

1. Precolumns require fused silica capillary tubing (360 μm outer diameter (o.d.) × 75 μm inner diameter (i.d.)), C18 reverse-phase resin (5–20 μm irregular diameter, 120 Å pore size), Kasil 1624 potassium silicate solution, formamide, and Teflon heat shrink tubing (0.06 o.d. × 0.012 i.d).
2. Analytical columns require fused silica capillary tubing (360 μm o.d. × 50 μm i.d.), C18 reverse-phase resin (5 μm diameter, 120 Å pore size), microcapillary laser puller with fused silica adapter, and polytetrafluoroethylene (PTFE) tubing (0.06 o.d. × 0.012 i.d.).
3. High-pressure column packer and sample loader (helium pressure bomb) (*see* **Note 1**).
4. CentriVap centrifugal vacuum concentrator.
5. Ultracentrifuge.
6. 5 μL glass capillary pipets.
7. Ultrasonic bath (e.g., the Branson 1200).
8. HPLC with nanoflow capabilities [14].
9. Orbitrap Fusion (Lumos) Tribrid Mass Spectrometer (Thermo Fisher Scientific) (*see* **Note 2**).

2.3 Software Requirements

1. Xcalibur (Thermo Fisher Scientific).
2. Byonic (Protein Metrics).
3. Molecular Operating Environment 2014 (MOE, Chemical Computing Group) [15] (*see* **Note 3**).

3 Methods

3.1 Sample Preparation

1. MHC class I and II peptide isolation is extracted as described in [16, 17] and stored in 10% acetic acid.
2. Take the samples to dryness using a vacuum concentrator.
3. Reconstitute samples in 0.1 M acetic acid to 1×10^7 cell equivalents (CEq) per μL (*see* **Note 4**).
4. Remove 5 μL of the 1×10^7 CEq/μL, and add to 4.8 μL 0.1 M acetic acid in a new Eppendorf tube.
5. Spike 0.2 μL of 500 fmol/μL peptide standard solution into the sample vial. The sample is now ready to be loaded onto the HPLC precolumn.

3.2 HPLC Column Preparation (See Note 5)

1. Create a Kasil frit on the precolumn fused capillary by first cutting the 360 μm o.d. × 75 μm i.d. fused silica capillary tubing to approximately 15 cm. In a glass vial, add 300 μL Kasil to 100 μL formamide, and vortex quickly (*see* **Note 6**). Briefly tap the end of the capillary tubing into the Kasil solution. Remove

quickly and wipe excess solution with a lint-free tissue. Bake overnight in an incubator set at 60 °C. After baking, cut the fused silica so that 1–2 mm of the Kasil frit remains at the end of the capillary.

2. Pack the precolumn by first making a slurry of C18 precolumn resin (5–20 μm irregular diameter, 120 Å pore size resin) in 80:20 (v/v) acetonitrile:isopropanol. Use the pressure bomb to pack 6–8 cm of the resin into the fritted fused silica capillary [18]. Rinse the column with several column volumes of 0.1 M acetic acid.
3. To make an analytical column, start by creating a laser-pulled restriction and bottleneck frit on the analytical column. Cut the 360 μm o.d. × 50 μm i.d. fused silica capillary tubing to approximately 15 cm. Use a flame to remove approximately 5 cm of the polyimide coating approximately 4 cm from the end of the capillary. Generate a 15 μm i.d. bottleneck using a laser-based micropipette puller. Then, make a slurry of C18 precolumn resin (5–20 μm irregular diameter, 120 Å pore size resin) in 80:20 (v/v) acetonitrile:isopropanol. Use the pressure bomb to pack the bottleneck frit with 2 mm of the precolumn resin. Wash the column with several volumes of 0.1 M acetic acid (*see* **Note 7**).
4. To pack the analytical column, make a slurry of C18 analytical column resin (5 μm diameter, 120 Å pore size resin) in 80:20 (v/v) acetonitrile:isopropanol. Use the pressure bomb to pack the analytical column with 6–8 cm of the analytical resin. Wash the column with several volumes of 0.1 M acetic acid. Using the laser-based micropipette puller, create a 2 μm emitter tip approximately 5 mm from the bottleneck frit.

3.3 Sample Loading

1. Using the pressure bomb, load the sample onto the precolumn at approximately 1 μL/min. Measure the flow rate using 5 μL glass capillary pipets.
2. Rinse the precolumn with ten column volumes of 0.1 M acetic acid.
3. Place an empty tube in the pressure bomb, insert the precolumn, and turn on the helium until air bubbles can be seen from the top of the precolumn. This serves to dry the precolumn, which prevents packing material from shifting.
4. Connect the precolumn to the analytical column using Teflon tubing.
5. Rinse the column on the HPLC until air bubbles are no longer visible in the emitter tip.

3.4 Equipment Setup

1. On the HPLC, set up a reverse-phase gradient using HPLC buffers A and B (e.g., 0–60 min, 0–60% B; 60–70 min,

60–100% B; 70–75 min, 100% B; 75–80 min, 100–0% B; 80–100, 0% B). Measure the flow rate with a 5 μL glass capillary tube, and adjust the flow rate so that the solvent comes out of the emitter tip at 60 nL/min. On the split flow system, a back pressure of 20–35 Bar at a flow rate of 0.2 mL/min is expected.

2. On the mass spectrometer, prepare an instrument method with the following parameters (*see* **Note 8**): MS1 resolution of 60,000 at 400 *m/z*, scan range of 300–1500 *m/z*, acceptable charge states of 2–4 (class I peptides) or 2–6 (class II peptides), and peptide monoisotopic precursor selection. Dynamic exclusion parameters are as follows: exclude after three times, if it occurs within 10 s, with a duration of 10 s and with a mass tolerance of ±10 ppm. For fragmentation spectra, use top-speed higher-energy collision dissociation (HCD), $n = 3$ s, an isolation width of 2 *m/z*, 25% collision energy, and Orbitrap detection (resolution at 30,000 at 400 *m/z*). The first mass should be detected at 100 *m/z*, precursor automatic gain control (AGC) target of 1×10^4, and maximum injection time of 100 ms. Use the product ion trigger tool to trigger electron transfer dissociation (ETD) on a precursor only if three of five *N*-acetylhexosamine (HexNAc; either N- or O-GlcNAc or O-linked α-*N*-acetylgalactosamine (O-GalNAc); *see* **Note 9**) fingerprint ions are present in an HCD spectrum at >5% relative abundance and ±0.1 Da. The ion masses are 204.0867, 186.0762, 168.0656, 144.0655, 138.0549, and 126.055. If sialic acid residues are expected (MHC class II), then the sialic acid fingerprint ions should be added to this list (292.1027, 274.0921) (*see* **Note 10**). In this case, ETD will be triggered if three of the seven ions are present at >5% relative abundance and ±0.1 Da. ETD parameters should be set as follows: isolation width of 2 *m/z*, 50 ms reaction time (*see* **Note 11**), and ETD reagent target 2×10^5, with ion trap (IT) detection. Set the first mass at 120 *m/z*, precursor AGC target 1e4, and maximum injection time of 100 ms.

3.5 Database Search of MHC-Associated Class I Glycopeptides

1. Set up a Byonic search by first downloading the human proteome database from websites such as UniProt or NCBI, and save it as a (.fasta) file (*see* **Note 12**). In the digestion and instrument parameters tab, use the following parameters: no cleavage sites, non-specific digestion specificity, precursor mass tolerance of 10 ppm, "both HCD and ETD" fragmentation type, fragment mass tolerance of 0.1 Da for high-resolution HCD, and/or 0.35 for low-resolution HCD and ETD. In the modifications tab, set the total common max to 3 and the total rare max to 1.

Add variable modifications of "Oxidation/+15.9949 @ M, common 1," "Deamidation/+0.9840 @ N, common 1," and "HexNAc/+203.0794 @ S,T, common 1" (*see* **Note 13**). Under the advanced tab, set the unassigned charge states to 2 and the maximum precursor mass to 5000, and compute precursor and charge assignments from MS1 and the maximum number of 2 precursors per MS2.

2. Once the search finishes, open the Excel output file from Byonic, and sort the results by "modification types," and move all "[+203]" (HexNAc) containing peptides to a new sheet in the excel file. Sort the new sheet by "[Log Prob]" (*see* **Note 14**).

3.6 Database Search Verification of MHC-Associated Class I Glycopeptides

1. First, confirm the residues in binding positions fit expectations. Variability between the different alleles of MHC class I α-chain results in defined peptide-binding motifs for each allele [19]. If the alleles have been defined for the cell line/tumor sample, both the P2 and P9 positions will have known characteristics (i.e., hydrophobic, basic, etc.). For instance, a sample containing only HLA-A2 peptides will have hydrophobic residues in P2 and P9. This serves as a way to weed out false positives.
2. Then, confirm that the peptide is glycosylated. If the peptide is truly glycosylated, the peptide's HCD spectra will have the HexNAc fingerprint (Fig. 1). Using Xcalibur Qual Browser, confirm this pattern is present in the low scan range of the HCD spectrum (the scan number will be found in column V).
3. Next, confirm the peptide sequence is correct by manually verifying the HCD and/or ETD spectra of the peptide. Peptide fragment ions can be calculated using "fragment ion calculators" readily available online (e.g., http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html). HCD fragment (b/y) ions will not be modified with a HexNAc, whereas the ETD fragment (c/z) ions will be modified.
4. Finally, confirm the site of modification with ETD. The site of modification will have drastic effects on binding and/or interaction with CD8+ T-cells, so localizing the site of modification is important [2–5]. ETD should be used to site localize the modification since ETD does not rely on collision for dissociation and the modification will remain attached to the modified residue [20]. To confidently assign a site of modification, both c and z ions surrounding the modification should be detected (Fig. 2). Alternatively, c or z ions can be used if all other possible sites of modification (N/S/T) have been assigned as unmodified (*see* **Note 15**).

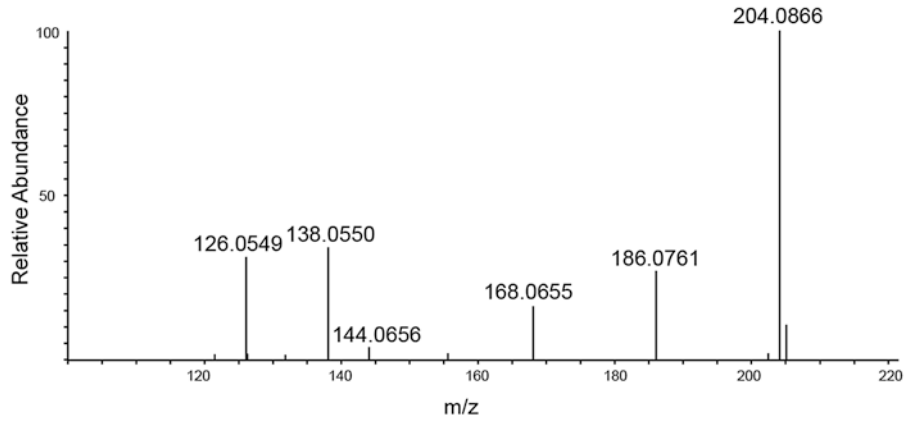


Fig. 1 Example HexNAc fingerprint. High-resolution HCD fragmentation of a HexNAc will result in these six low-mass ions ±10 ppm: 204.0867, 186.0762, 168.0656, 144.0655, 138.0549, and 126.055. If a sialic acid residue is present in the glycan, two additional ions will be found in the low-mass region ±10 ppm: 292.1027 and 274.0921. These ions can confirm the presence of a glycan on a peptide

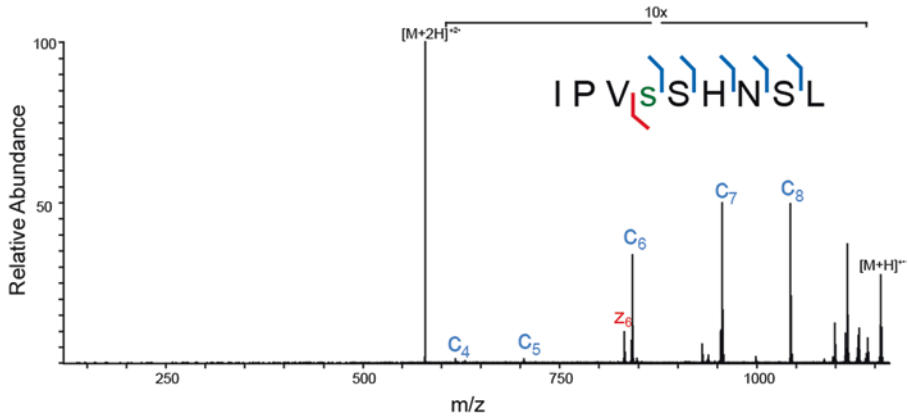


Fig. 2 Example ETD spectrum of an MHC-associated class I glycopeptide. The peptide, IPVsSHNSL, is modified by an O-GlcNAc at the S4 position. Fragment ions that define the amino acid sequence are labeled as c (blue) and z (red). "10x" indicates that the relative abundance has been magnified by ten times

3.7 Database Search of MHC-Associated Class II Glycopeptides

1. Byonic search parameters are the same as class I (*see* Subheading 3.5), except for the following parameters. Under the modifications tab, set the total common max to 3, total rare max to 1, and variable modifications of "Oxidation/+15.9949 @ M, common 1" and "Deamidated/+0.9840 @ N, common 1." Then, under the glycans tab (*see* **Note 16**), select "enter/edit," and add both the "O-glycan 6 most common" database (common 1) and the "N-glycan 57 human plasma" database (common 1). Under the advanced tab, apply charge state 3 to unassigned ions; set the maximum precursor mass to 10,000.
2. After the search is completed, open the Excel output file from Byonic, and sort the results by "Glycans." Move all putative

glycan-containing peptides to a new sheet in the Excel file. Sort this sheet by "[Log Prob]" (*see* **Note 14**).

3.8 Database Search Validation of MHC-Associated Class II Glycopeptides

1. First, confirm the peptide is glycosylated (*see* Subheading 3.6, **step 2**).
2. Next, confirm the glycan structure is correct (*see* **Note 17**). Byonic will assign a structure for each site of modification. To confirm this is correct, open the HCD spectrum, and observe the neutral losses. If the sequential losses are doubly charged, calculate the $[M + 2H]^{+2}$ mass from the accurate mass; if the losses are triply charged, calculate the $[M + 3H]^{+3}$ mass. From the corresponding mass, follow the neutral losses to obtain the intact glycan structure. Keep in mind that the m/z of neutral losses should be multiplied by z (charge state) to obtain the mass of the monosaccharide structure. *See* Fig. 3 for an example of an HCD spectrum of a glycopeptide with a high-mannose structure.
3. Confirm the peptide sequence is correct (*see* Subheading 3.6, **step 3**).
4. Confirm the site of modification with ETD (*see* Subheading 3.6, **step 4**).

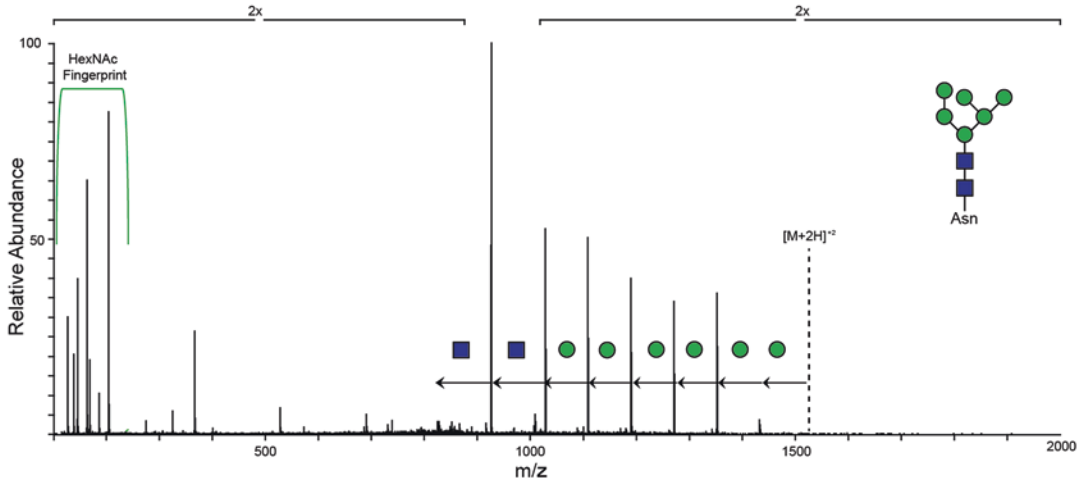


Fig. 3 Example HCD spectrum of a MHC-associated class II glycopeptide. The precursor mass is triply charged at mass of 1009.4465, making the $[M + 2H]^{+2}$ equal to 1513.6661 (denoted by the dotted line; it is common to deplete the intact glycopeptide using HCD). From the $[M + 2H]^{+2}$ mass, double-charged neutral losses are observed corresponding to 6 hexose (mannose) and 2 HexNAc (in this case, GlcNAc) residues. As N-terminal glycans have the core $HexNAc_2$-$hexose_3$ structure at the base of the glycan, we can conclude this is a high-mannose N-linked glycan. A putative structure is shown at right. "2x" indicates that the relative abundance has been magnified by two times

3.9 Manual Sequencing of MHC-Associated Glycopeptides

1. Open the (.raw) file in Xcalibur Qual Browser.
2. Select the chromatogram in the right-hand corner (pin).
3. Right click and select "Ranges" in the drop-down menu.
4. In the scan filter, type "FTMS + c NSI r d Full ms2".
5. In the plot type, select "Mass Range".
6. In the ranges, type "204.0867".
7. Under the automatic processing tab, under mass tolerance, check "Use user defined," and enter 10 ppm for high-resolution (Orbitrap) data. For low-resolution (IT) data, widen this to 500 ppm and click apply. This will plot all HCD spectra that contain the accurate mass for a protonated HexNAc ion (*see* **Note 18**). Confirm that the according spectrum contains the six HexNAc fingerprint ions.
8. Find the corresponding ETD spectrum (or spectra) for the peptide.
9. Use both the HCD and ETD spectra to sequence the peptide and the ETD spectra to site localize the glycopeptides. Tutorials for de novo HCD sequencing [21] and ETD sequencing [22] are available elsewhere (*see* **Note 19**).

3.10 Preparation and Overlay of X-Ray Crystal Structures

1. From the Protein Data Bank (PDB), download structural information (PDB IDs: 4OZG, 3T0E, 4C56, 4E41, and 1FYT) [23–27] for the five crystallized MHC class II-peptide-TCR (T-cell receptor) ternary complexes as (.pdb) files.
2. Open structures in MOE using "*File→Open*" and selecting each (.pdb) file (*see* **Note 20**).
3. Add the unresolved hydrogen atoms to each species. In MOE, this is done using "*Compute→Prepare→Protonate 3D*" with the standard parameters (*see* **Note 21**).
4. Align the sequences of the proteins of the different crystal structures. In MOE, this is achieved using "*Protein→Align/Superpose*" and then choosing "*Align*" with "*Sequence Only*".
5. Without affecting any bonds, overlay the different complexes. In MOE, this is best accomplished by selecting all atoms of an MHC-peptide-TCR ternary complex and then manually repositioning the individual complexes using a combination of "*Shift*", "*Alt*", and the mouse wheel. Once crudely overlaid, select the residues of all MHC α subunits and use "*Protein→Align/Superpose*" and then choosing "*Superpose*" with "*Selected Residues*" and "*Structure Only*" (*see* **Note 22**).
6. Save the overlaid complexes as a (.moe) file for later use.

3.11 Construction of Glycopeptides

1. Create a copy of the crystallographic peptide ligand GELIGILNAAKVPAD (PDB ID: 4E41) from the overlaid MHC-peptide-TCR ternary complex (*see* **Note 23**). In MOE, this is achieved by double-clicking on the peptide and then performing "*Edit→Copy*" and "*Edit→Paste*".
2. Adjust the copied peptide until the desired sequence is obtained. In MOE, this is done using "*Protein→Protein Builder*" to delete, mutate, or append individual residues. While adjusting the sequence, also adjust the side chains of affected residues to ensure they are built in their most stable conformation. In MOE, this is done using "*Protein→Protein Builder*" and analyzing the different conformations and energies of different conformers (*see* **Note 24**).
3. Similar to before, create copies of the crystallographic glycans-$GlcNAc_2Mannose_3$ of N478 and glycans-$GlcNAc_2Fucose_1Mannose_3$ of N243 of gp120 (PDB ID: 2BF1) [28] by selecting all atoms of each glycan and duplicating these species by performing "*Edit→Copy*" and "*Edit→Paste*" (*see* **Note 25**).
4. As before, add unresolved hydrogen atoms to the two glycan fragments. In MOE, use "*Compute→Prepare→Protonate 3D*" with the standard parameters.
5. Without affecting any bonds, move the desired glycan fragment(s) so that the anomeric carbon of the first GlcNAc moiety is properly positioned adjacent to the desired asparagine side chain. In MOE, this is done by selecting all atoms of the glycan and then using a combination of "*Shift*", "*Alt*", and the mouse wheel to reposition the species.
6. Attach the glycan to the asparagine side chain, ensuring that the new glycosidic linkage is β and that the resulting amide is in the *trans* arrangement. In MOE, select the anomeric carbon and amide nitrogen. Then use "*Edit→Build→Molecule*" to create the new single bond.
7. Minimize the modified asparagine residue to ensure the species is unstrained. During minimization, the asparagine residue and glycan should be tethered, and all remaining atoms of the glycopeptide should be held fixed to prevent larger peptide rearrangement. In MOE, select the appropriate atoms, and use "*Constrain→Tether*" or "*Constrain→Fix*," respectively. Following this, minimize the entire glycopeptide using "*Protein→Protein Builder*" in the "Amber 10:EHT" force field [29].
8. Save the glycopeptide as a (.moe) file before further use.

3.12 Modeling of Glycopeptides

1. Ensure that MOE is using the "Amber 10:EHT" force field and that the desired glycopeptide is the only species that is open and active.
2. Adjust the glycopeptide so that the atoms of the glycopeptide are held fixed except those of the modified asparagine residue and glycan. In MOE, this is done using "*Constrain→Fix*" or "*Constrain→Free*" similar to before.
3. Perform a conformational search to create a database of glycopeptide conformers that vary only in the arrangement of their glycans. In MOE, this is accomplished using "*Compute→Conformations→Search*" (*see* **Note 26**); this method requires some alterations to the standard parameters to improve results: Method (LowModeMD), Rejection Limit (1000), Iteration Limit (30,000), MM Iteration Limit (5), RMSD Limit (0.01), Energy Window (50), Conformation Limit (30,000), Exclude fixed atoms from shape descriptors (On).
4. Once the conformational search is complete, analyze the results to determine how the glycopeptide may interact with TCRs. In MOE, it is easiest to perform this analysis by (1) reopening the overlaid MHC-peptide-TCR ternary complexes from Subheading 3.7, **step 6**, (2) opening the database (.mdb) of glycopeptide conformers created in Subheading 3.9, **step 3**, and (3) scanning through the results using "*File→Browse*".

4 Notes

1. The helium pressure bombs used in these procedures were prepared in-house; however, Proxeon offers a similar system commercially.
2. Any mass spectrometer equipped with ETD is technically capable of performing the procedures outlined here. However, due to the increased analytical capabilities of the Tribrid series of mass spectrometers, the likelihood of detecting glycosylated species is much higher with these instruments. If these are not available, an enrichment method prior to mass spectrometric analysis is recommended [13].
3. Several programs can be used to perform the molecular modeling described in this study. We opted to use the Molecular Operating Environment software suite, which is available for purchase from the Chemical Computing Group. This software package provides an intuitive interface that allows for visualizing, building, aligning, and simulating molecules using a desktop computer. Whenever possible, we have tried to include

both general and MOE-specific instructions to aid readers looking to perform similar work.

4. For cell lines, cell equivalents are simply the number of cells that were used to extract the MHC-associated peptides. For tissue samples, 1 g of tumor/normal tissue (prior to homogenization/lysis) is assumed to be 1×10^9 cell equivalents.
5. This setup was used for all of the authors' experiments; however, several nanoflow HPLC instruments are now available commercially. The most important aspect of the setup is the low flow rate, as it is inversely proportional to sensitivity and will increase the likelihood of detecting low abundance analytes.
6. Polymer will begin to form in the vial if the solution is not vortexed quickly enough. If this is the case, dispose of the polymer-containing solution and start again.
7. The bottleneck frit can be easily clogged with the irregular packing solution. If the column stops flowing at >500 psi during the 0.1 M acetic acid rinse, it is necessary to make a new frit.
8. The instrument parameters listed are for the Fusion instrument and Xcalibur software (version 3.0). Similar methods can be made on the older version of Xcalibur (on instruments such as the Orbitrap, Velos, or Elite, where ETD is available) but are not detailed here.
9. As O-GlcNAc and O-GalNAc are isobaric, HexNAc is used when the instrument, search algorithm, or some other experimental condition(s) cannot inherently distinguish between the two modifications.
10. Ion trap detection is also an option after HCD fragmentation as it is a more sensitive detection technique. However, because it is low resolution, the masses for the ETD trigger should have a wider tolerance, such as ±0.35 Da.
11. Alternatively, ETD can be set to "calibrated reaction times" which uses various ETD reaction times based on the charge state of the precursor ions.
12. These parameters can be applied to other mass spectrometry software suites (SEQUEST, OMSSA, Mascot, etc.), but Byonic is currently the best search algorithm for glycoproteomics. That being said, all search algorithms struggle with "no enzyme" searches since the search space is so large.
13. As per [13], N-linked GlcNAc and (N/S/T) + GlcNAc-hexose peptides have been isolated from MHC class I samples. These modifications can be added into Byonic as well but will add additional time to the search. To add N-linked HexNAc, simply check the "N" box when adding S and T HexNAc. To add

the latter, select "Hex1HexNAc1" from the modification drop-down, and add N, S, and/or T. In fine control, select Variable—rare 1.

14. Anything under a [Log Prob] value of 5 is unlikely to be correct. The authors stop validating at this point in the analysis.
15. MHC class I peptides are short, eight to ten amino acid peptides, with few sites of positive charge. Thus, the ETD spectra generated will likely be from doubly-charged precursors and may not have enough sequence coverage to localize the site of glycosylation. If this is the case, check for an ETD spectrum of the $[M + 3H]^{+3}$ mass, which can potentially yield better ETD spectra.
16. Any of the glycan databases can be used for these searches, but the larger the glycan database, the longer the search time. Alternatively, a user can input specific glycans, which may be useful, especially if glycomics data has been acquired on a specific sample.
17. This strategy works best for N-linked glycans since they generally follow three defined templates: high-mannose, hybrid, and complex glycans. O-linked structures do not follow similar templates and thus fragment in more complex ways. These may not be interpretable in the same manner.
18. Hundreds (or potentially thousands) of spectra will have the HexNAc ion present; however, not all spectra are good enough to manually sequence. Those having Orbitrap ion counts $<1 \times 10^5$ are likely too weak to be sequenced.
19. If the precursor mass is over 1000 m/z, the charge density of the peptide can be too low to obtain fruitful ETD spectra. In this case, look for higher charge states (+3, +4, etc.) of the precursor. If available, this will likely provide more fragmentation information and should allow for sequencing of the peptide and site localization of the modification. An example of a +3 (gt;1000 m/z) vs a + 4 (lt;1000 m/z) is shown in Fig. 4.
20. When the files are opened, MOE will load a prompt with several parameters that can be changed. In this case, it is best to use the standard parameters.
21. For in-depth modeling studies (i.e., simulations), more structure preparation will be required to account for (a) unresolved atoms and (b) solvation of the complex. In this case, the unresolved atoms of the crystal structures are primarily found in the TCR α and β subunits, and they are distant from the peptide-binding groove. Thus, any alterations are speculative and unnecessary for this study.
22. There are various methods and weighting factors that can be used to perform the superposition. Because these glycopep-

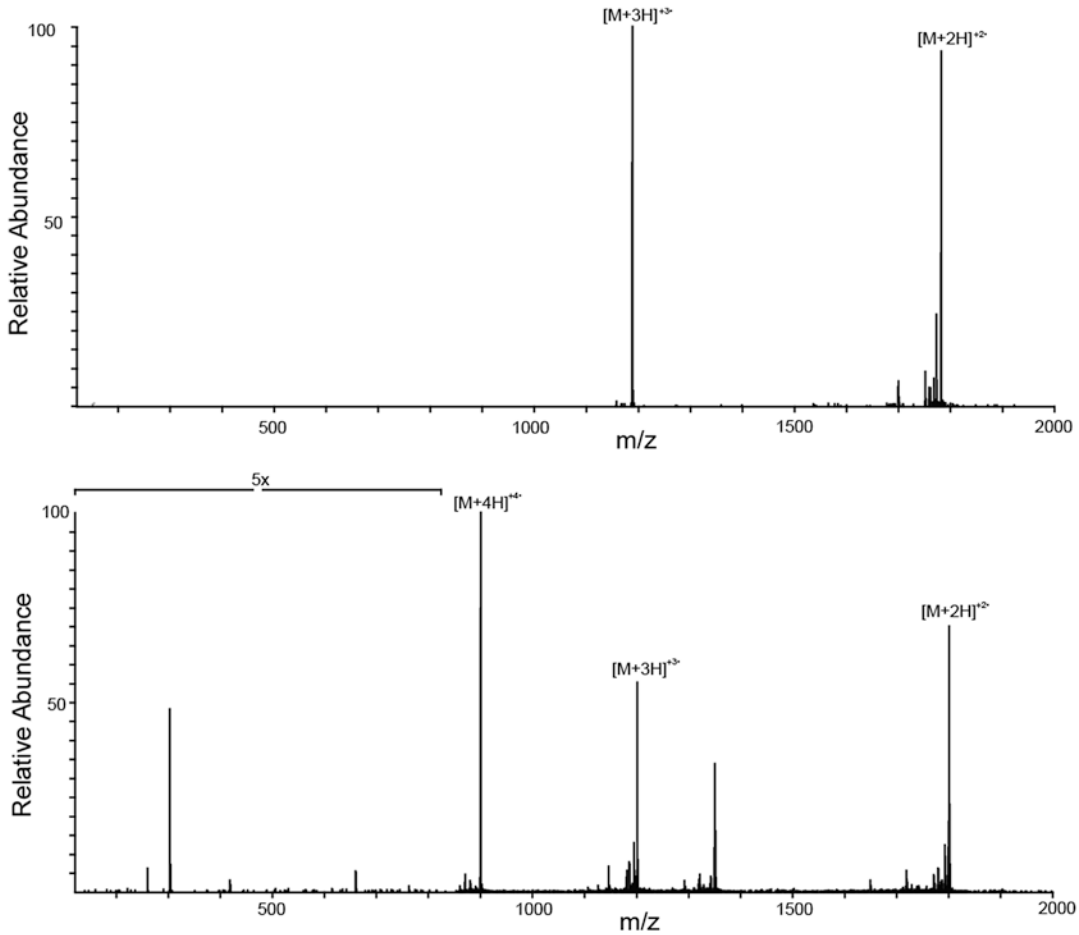


Fig. 4 $[M + 3H]^{+3}$ vs $[M + 4H]^{+4}$ ETD spectra of an MHC-associated class II glycopeptide. In the +3 spectrum (top panel), the charge density of the glycopeptide is not high enough to allow for detection of fragment ions and is instead dominated by charge-reduced species. In the +4 spectrum (bottom panel), the increased charge density improves the fragmentation which allows for peptide sequencing and glycan site localization. "5x" indicates that the relative abundance has been magnified by five times

tides were shown to bind to MHCs experimentally, we opted to emphasize the similarity in the presentation of the glycopeptide-MHC complex to the TCRs.

23. The peptide portion of each glycopeptide was modeled using coordinates from the X-ray crystal structure of the peptide GELIGILNAAKVPAD (PDB ID: 4E41), as it has the most resolved residues of the peptide ligands in any of the X-ray crystal structures of the MHC-peptide-TCR ternary complexes.
24. It is essential that the atoms of the MHC α and β subunits—but not the crystallographic peptide ligand or TCR α and β subunits—are held fixed (by selecting the atoms and using

"*Constrain→Fix*") and are active (by selecting the atoms and using "*Edit→Potential→Active*") when calculating the energy of the side chain conformers. If this is not done, then the program will determine each conformer's energy in a vacuum rather than in the unique steric and electronic environment of the MHC peptide-binding groove.

25. The carbohydrate portion of the glycopeptide was modeled using coordinates from the X-ray crystal structure of the glycans-$GlcNAc_2Mannose_3$ on N478 and glycans-$GlcNAc_2Fucose_1Mannose_3$ on N243 of gp120 (PDB ID: 2BF1), as these glycans are well resolved and lie on solvent-exposed portions of gp120. The chosen glycans are smaller than those detected by mass spectrometry, but they are, to our knowledge, the best-resolved structures of these large glycan fragments.
26. There are many different ways of generating a library of glycopeptide conformers. The "*LowModeMD*" option used by the "*Conformational Search*" is essentially a coarse molecular dynamics simulation. Here, the simulation is parameterized so that it performs 30,000 total iterations of random atom movement, saving the unique conformers (heavy atom RMSD >0.01) as separate entries in the database. Depending upon the species, our conformational searches yielded between a few hundred and a few thousand different conformers.

Acknowledgments

The authors would like to acknowledge Jane V. Aldrich (U Florida), Dina L. Bai, Jeffrey Shabanowitz, and Donald F. Hunt (U Virginia) for their technical and financial support. This work was supported by a research grant from the Melanoma Research Alliance and by an NIH grant AI033993 to D.F.H. S.A.M. is currently funded by an NIH F32 Postdoctoral fellowship.

References

1. Pinho SS, Reis CA (2015) Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer 15:540–555
2. Haurum JS, Tan L, Arsequell G, Frodsham P, Lellouch AC, Moss PA et al (1995) Peptide anchor residue glycosylation: effect on class I major histocompatibility complex binding and cytotoxic T lymphocyte recognition. Eur J Immunol 25:3270–3276
3. Haurum JS, Arsequell G, Lellouch AC, Wong SY, Dwek RA, McMichael AJ et al (1994) Recognition of carbohydrate by major histocompatibility complex class I restricted, glycopeptide-specific cytotoxic T lymphocytes. J Exp Med 180:739–744
4. Kastrup IB, Stevanovic S, Arsequell G, Valencia G, Zeuthen J, Rammensee HG et al (2000) Lectin purified human class I MHC-derived peptides: evidence for presentation of glycopeptides in vivo. Tissue Antigens 56:129–135
5. Glithero A, Tormo J, Harurum JS, Arsequell G, Valencia G, Edwards J, Springer S et al

(1999) Crystal structures of two H-2Db/glycopeptide complexes suggest a molecular basis for CTL cross-reactivity. Immunity 10:63–74

6. Dzhambazov B, Holmdahl M, Yamada H, Lu S, Vestberg M, Holm B, Johnell O, Kilberg J, Holmdahl R (2005) The major T cell epitope on type II collagen is glycosylated in normal cartilage but modified by arthritis in both rats and humans. Eur J Immunol 35(2):357–366
7. Mastrangelo A, Colasanti T, Barbati C, Pecani A, Sabatinelli D, Truglia S, Massaro L et al (2015) The role of posttranslational protein modifications in rheumatological diseases: focus on rheumatoid arthritis. J Immunol Res 2015:712490
8. Housseau F, Moorthy A, Langer DA, Robbins PF, Gonzales MI, Topalian SL (2001) N-linked carbohydrates in tyrosinase are required for its recognition by human MHC class II-restricted CD4+ T cells. Eur J Immunol 31:2690–2701
9. Chicz RM, Urban RG, Gorga JC, Vignali DA, Lane WS, Strominger JL (1993) Specificity and promiscuity among naturally process peptides bound to HLA-DR alleles. J Exp Med 178(1):27–47
10. Dengjel J, Rammensee HG, Stavnovic S (2005) Glycan side chains on naturally presented MHC class II ligands. J Mass Spectrom 40(1):100–104
11. Bern M, Kil YJ, Becker C (2012) Byonic: advanced peptide and protein identification software. Curr Protoc Bioinformatics 13:Unit 13.20
12. Malaker SA, Ferracane MJ, Depntieu FR, Zarling AL, Shabanowitz J, Bai DL, Topalian SL, Engelhard VH, Hunt DH (2017) Identification and characterization of complex glycosylated peptides presented by the MHC class II processing pathway in melanoma. J Prot Res 16:228–237
13. Malaker SA, Penny SA, Steadman LG, Myers PT, Loke JC, Raghavan M, Bai DL, Shabanowitz J, Hunt DF, Cobbold M (2017) Identification of glycopeptides as posttranslationally modified neoantigens in leukemia. Cancer Immunol Res 5(5):376–384
14. Martin SE, Shabanowitz J, Hunt DF, Marto JA (2000) Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 72:4266–4274
15. Molecular Operating Environment (MOE), 2014.09; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015
16. Cobbold M, De La Pena H, Norris A, Polefrone JM, Qian J, English AM, Abelin JG, Malaker SA et al (2013) MHC Class I–Associated Phosphopeptides Are the Targets of Memory-like Immunity in Leukemia. Sci Transl Med 5(203):203ra125
17. Depontieu FR, Qian J, Zarling AL, McMiller TL, Salay TM, Norris A, English AM et al (2009) Identification of tumor-association, MHC class I-restricted phosphopeptides as targets for immunotherapy. Proc Natl Acad Sci U S A 106(29):12073–12078
18. Udeshi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL (2008) Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. Nat Protoc 3:1709–1717
19. Engelhard VH, Brickner AG, Zarling AL (2002) Insights into antigen processing gained by direct analysis of the naturally processed class I MHC associated peptide repertoire. Mol Immunol 39(3–4):127–137
20. Syka JE, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A 101(26):9528–9533
21. Papayannopooulos IA (1995) The interpretation of collision-induced dissociation of tandem mass spectra of peptides. Mass Spectrom Rev 14(1):49–73
22. Hunt DF, Shabanowitz DF, Bai DL (2015) Peptide sequence analysis by electron transfer dissociation mass spectrometry: a web-based tutorial. J Am Soc Mass Spectrom 26(7):1252–1255
23. Petersen K, Montserrat V, Mujico JR, Loh KL, Beringer DX, van Lummel M et al (2014) T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. Nat Struct Mol Biol 21(5):480–488
24. Yin Y, Wang XX, Mariuzza RA (2012) Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. Proc Natl Acad Sci U S A 109(14):5405–5410
25. Rodstrom KE, Elbing K, Lindkvist-Petersson K (2014) Structure of the superantigen staphylococcal enterotoxin B in complex with TCR and peptide-MHC demonstrates absence of TCR-peptide contacts. J Immunol 193(4):1998–2004
26. Deng L, Langley RJ, Brown PH, Xu G, Teng L, Wang Q et al (2007) Structural basis for the recognition of mutant self by a tumor-specific, MHC class II-restricted T cell receptor. Nat Immunol 8(4):398–408

27. Hennecke J, Carfi A, Wiley DC (2000) Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. EMBO J 19(21):5611–5624
28. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. Nature 433(7028):834–841
29. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner PA (1984) New force-field for molecular mechanical simulation of nucleic-acids and proteins. J Am Chem Soc 106(3):765–784



Chapter 18

Genome-Based Bioinformatic Prediction of Major Histocompatibility (MHC)

Simon J. Foote

Abstract

Over the last 17 years, a large amount of knowledge has been accumulated on various aspects of major histocompatibility complex (MHC) molecules. In conjunction, numerous algorithms and tools have been developed to screen protein molecules for these MHC receptor sites. By combining these computational tools and databases with genomic sequence information that is now widely available for a vast range of organisms, it is possible to screen whole genomes for MHC epitopes. By pre-screening these genomes, it allows the researcher to narrow down possible protein targets for further analysis by traditional tools such as gene knockouts and animal efficacy studies.

Key words MHC epitope prediction, MHC ligand, T cell, Antigen

1 Introduction

Antigenic epitopes play a key role in eliciting a T-cell response against intracellular and extracellular pathogens including bacteria, viruses, and parasites. T cells contain two receptors, a T-cell receptor (TCR) which recognizes antigens and a co-receptor which binds to major histocompatibility complex (MHC) molecules. There are two main types of T cells, cytotoxic T cells and helper T cells. The cytotoxic T cell's co-receptor (CD8) recognizes an MHC class I antigen complex and in combination with the TCR binding to the antigen leads to T-cell activation and eventual cell death. Helper T cells control the immune response by directing other cells to fight infected cells. The helper T cell's co-receptor (CD4) recognizes an MHC class II antigen complex which activates the T cell to release cytokines that activate other cell types such as B cells. The antigen presented on the MHC molecules is typically a peptide fragment, and the resulting complex is referred to as a T-cell epitope. The antigen peptides are generated via one of two pathways depending upon whether the pathogen is intra- or extracellular. In both pathways, specific proteases are responsible

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_18, © Springer Science+Business Media, LLC, part of Springer Nature 2019

for cleaving the proteins into smaller peptides which then bind to MHC class I or class II molecules. MHC class I molecules accommodate peptides 8–10 amino acids in length and usually contain an allele-specific motif sequence with conserved N-terminal and C-terminal residues [1]. MHC class II molecules bind peptides 11–25 amino acids in length. They are more variable in size but usually contain a 9–10 residue core region [2]. By using this gathered epitope information, multiple algorithms, software tools, and databases have been developed to search for MHC ligands.

The aim of this chapter will be to describe a selection of these tools and how they can be used together to search for MHC epitopes in whole genomes of bacteria. As an example of this genome-based analysis, the genome of *Francisella tularensis* subsp. *tularensis* (strain SCHU S4) which leads to typhoidal tularemia upon inhalation will be analyzed for probable peptides that could bind to MHC class I or class II molecules. It is known that SCHU S4 express antigens unique to subsp. *tularensis* and these might contribute to protective immunity [3]. The research involved the use of deletion mutants of SCHU S4 which showed in mice that a live vaccine strain (LVS) elicited immunity to challenge with subsp. *tularensis* is known to be dependent on the actions of interferon gamma (IFNγ) and CD4+ and CD8+ T cells. Either CD4+ T cells or CD8+ T cells appear to be sufficient for survival, as mice deficient in either T-cell subset are capable of resolving both primary and secondary LVS infections. Although this has been shown, the actual epitopes that induce this MHC response have not been identified.

2 Materials

For additional software and databases, *see* Tables 1 and 2.

2.1 MHC Prediction Software

1. Rankpep [9]—predicts peptide binders to MHC I and MHC II molecules from protein sequence/s or sequence alignments using Position-Specific Scoring Matrices (PSSMs) and webform (multi-sequence input), http://imed.med.ucm.es/Tools/rankpep.html.
2. IEDB Analysis Resource v2.18.1 [10]—multi-algorithmic depending upon chosen alleles (matrices, artificial neural network, average relative binding), MHC I or MHC II, webform, and local installation options, http://tools.iedb.org/main/tcell.

2.2 Databases

1. IEDB—Immune Epitope Database (IEDB) is a database containing curated immunological data for antibody and B- and T-cell epitopes derived from humans, nonhuman primates, rodents, and other animal species, http://www.iedb.org.

Table 1
A selection of alternative software tools for MHC prediction (see Note 1)

Name	Description	Link
ProPred-I [4]	Matrix-based, employs matrices from 47 MHC I alleles, webform (single sequence)	http://www.imtech.res.in/raghava/propred1
NetMHCpan-4.0 [5]	Web server predicts binding to any known MHC I molecule using artificial neural networks, multi-sequence input, also downloadable version	http://www.cbs.dtu.dk/services/NetMHCpan
NetMHCIIpan-3.2 [6]	Web server predicts binding of peptides to more than 500 HLA-DR human MHC II alleles using artificial neural networks (ANNs)	http://www.cbs.dtu.dk/services/NetMHCIIpan
TEPITOPEpan [7]	MHC II matrix-based, human allele specific, webform	http://datamining-iip.fudan.edu.cn/service/TEPITOPEpan/index.html
SYFPEITHI	Matrix-based prediction of MHC I and MHC II epitopes, web-based, and downloadable versions	http://www.syfpeithi.de

Table 2
A selection of databases containing MHC ligand and peptide information

Name	Description	Link
MHCBN-4.0 [8]	Curated database consisting of detailed information about major histocompatibility complex (MHC) binding, non-binding peptides and T-cell epitopes	http://crdd.osdd.net/raghava/mhcbn/index.html
SYFPEITHI	Database of published MHC ligands and peptide motifs	http://www.syfpeithi.de

2.3 Protein Localization Software

1. PSortB v3.0.2 [11]—bacterial localization prediction, http://www.psort.org/psortb.
2. SignalP v4.1 [12]—predicts the presence and location of signal peptide cleavage sites, http://www.cbs.dtu.dk/services/SignalP.
3. SecretomeP v2.0 [13, 14]—predicts nonclassical protein secretion, http://www.cbs.dtu.dk/services/SecretomeP.
4. LipoP v1.0 [15]—predicts lipoproteins and discriminates between lipoprotein signal peptides, other signal peptides, and N-terminal membrane helices, http://www.cbs.dtu.dk/services/LipoP.

3 Methods

Whole genome-based analysis for MHC epitopes can be accomplished for any completed genome available from the National Center for Biotechnology Information (NCBI). Files containing the protein information for such genomes can be downloaded from NCBI via ftp. (ftp://ftp.ncbi.nlm.nih.gov/genomes). In this example, the *F. tularensis* subsp. *tularensis* strain SCHU S4 genome will be analyzed (NCBI RefSeq = NC_006570).

3.1 Data Gathering and Target Selection

1. Download protein sequences for *F. tularensis* subsp. *tularensis* strain SCHU S4 genome (1604 proteins) from NCBI, and save as a text file (ftp.ncbi.nih.gov/genomes/Bacteria/Francisella_tularensis_SCHU_S4_uid57589/NC_006570.faa) (*see* **Note 2**).
2. Identify MHC alleles to target (*see* **Note 3**).
   For MHC I, use the BALB/c mouse alleles: H-2-Db, H-2-Dd, H-Kb, and H-Kd.
   For MHC II, use the mouse alleles: H2-IAd and H2-IEd.

3.2 MHC I and MHC II Analysis with RankPep

1. FASTA files containing up to 100 proteins are used for input (17 files total) (*see* **Note 4**).
2. Set parameters on the webform (http://imed.med.ucm.es/Tools/rankpep.html) as indicated.
   (a) Select PSSM = chose alleles depending upon whether it is an MHC I or MHC II analysis.
   (b) Input type = FASTA sequences and upload sequences by choosing an input file from **step 1** above.
   (c) Allowed peptide lengths = all.
   (d) Binding threshold = 2%.
   (e) Proteasome cleavage = on.
3. Save results as an html file.
4. Convert the html file to text, and filter for top-scoring peptides by using a cutoff of IC50 nM $\leq$ 65 on the top ranking algorithm for MHC I and IC50 nM $\leq$ 50 for MHC II (*see* **Note 5**) (*see* Tables 3 and 4 for MHC I and MHC II results, respectively).
5. Extract the identified proteins into separate FASTA files for further analysis.

3.3 Protein Localization Analysis

For effective epitope presentation, it is important to verify whether top-scoring proteins from the above analyses are presented outside the cell, thus being accessible to MHC binding proteins.

Table 3
MHCI peptides identified by RankPep with binding thresholds ≤65 nM along with the corresponding protein localization information as identified by PSortB, SecretomeP, LipoP, and SignalP

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Binding threshold	Location	Secretion possible	Class	Signal peptide present
YP_169146.1	10mer_H2_Kd.p.mtx	TYIDGGKGVL	29.131	65.79	Cytoplasmic	No	CYT	No
YP_169174.1	10mer_H2_Kd.p.mtx	TYIGVGILLL	29.978	67.7	Cytoplasmic membrane	No	TMH	No
YP_169175.1	10mer_H2_Dd.p.mtx	AGGYINYFFL	35.566	67.01	Cytoplasmic membrane	No	TMH	No
YP_169197.1	9mer_H2_Kd.p.mtx	KYIKENTGL	27.098	66.34	Cytoplasmic	No	CYT	No
YP_169451.1	10mer_H2_Dd.p.mtx	SGPYSLLFIF	36.737	69.22	Cytoplasmic	No	CYT	No
YP_169464.1	8mer_H2_Kd.p.mtx	YYKNYQNL	31.016	72.2	Cytoplasmic	No	CYT	No
YP_169465.1	8mer_H2_Kd.p.mtx	RIPNYAKL	28.29	65.85	Cytoplasmic	No	CYT	No
YP_169519.1	**11mer_H2_Kd.p.mtx**	**LYQPNAVTQNK**	**31.314**	**68.02**	**Periplasmic**	**No**	**SpI**	**Yes**
YP_169540.1	10mer_H2_Kd.p.mtx	KYPSFVIQNI	29.332	66.24	Cytoplasmic membrane	No	TMH	No
YP_169618.1	9mer_H2_Kd.p.mtx	KYICIVCGL	30.341	74.28	Cytoplasmic	No	CYT	No
YP_169630.1	8mer_H2_Kd.p.mtx	YYQFYQDL	32.137	74.81	Cytoplasmic	No	CYT	No
YP_169635.1	9mer_H2_Kd.p.mtx	FYICNNTTL	30.375	74.36	Cytoplasmic membrane	No	TMH	No
YP_169699.1	8mer_H2_Kd.p.mtx	YYHTFTTL	28.129	65.48	Cytoplasmic membrane	No	CYT	No
YP_169751.1	**10mer_H2_Kd.p.mtx**	**SYVDFGITLL**	**29.636**	**66.93**	**Outer membrane**	**Yes**	**CYT**	**Yo**
YP_169784.1	9mer_H2_Dd.p.mtx	VGPKRRKAL	32.458	77.25	Cytoplasmic	No	CYT	No

(continued)

Table 3
(continued)

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Binding threshold	Location	Secretion possible	Class	Signal peptide present
YP_169855.1	9mer_H2_Kd.p.mtx	RYLCIPTLI	29.102	71.24	Cytoplasmic membrane	No	TMH	No
YP_169902.1	8mer_H2_Kd.p.mtx	YYIFYTVL	31.481	73.28	Cytoplasmic membrane	No	TMH	No
YP_169903.1	10mer_H2_Kd.p.mtx	PYPPFITSTL	29.545	66.72	Cytoplasmic	No	CYT	No
YP_169913.1	8mer_H2_Kd.p.mtx	YILFFEPL	30.741	71.56	Cytoplasmic	No	CYT	No
YP_169938.1	8mer_H2_Kd.p.mtx	YQPDFCLI	29.774	69.3	Cytoplasmic	No	CYT	No
YP_169943.1	9mer_H2_Kd.p.mtx	LYQQMPTSL	28.979	70.94	Cytoplasmic membrane	No	TMH	No
YP_169947.1	8mer_H2_Kd.p.mtx	YIYVFSGL	30.86	71.83	Cytoplasmic	No	CYT	No
YP_169966.1	8mer_H2_Kd.p.mtx	FYIVFFKL	29.282	68.16	Cytoplasmic membrane	No	CYT	No
YP_169988.1	9mer_H2_Kd.p.mtx	RYAKNNEEL	27.72	67.86	Cytoplasmic	No	CYT	No
YP_170040.1	9mer_H2_Kd.p.mtx	LYNQNIELV	26.913	65.89	Cytoplasmic	No	CYT	No
YP_170050.1	8mer_H2_Kd.p.mtx	YYIDFESL	34.041	79.24	Cytoplasmic	No	CYT	No
YP_170132.1	8mer_H2_Kd.p.mtx	YYISYKDI	28.071	65.34	Cytoplasmic	Yes	CYT	No
YP_170216.1	**9mer_H2_Kd.p.mtx**	**KYAQNQHKL**	**26.662**	**65.27**	**Outer membrane**	**Yes**	**SpII**	**Yes**
YP_170227.1	10mer_H2_Kd.p.mtx	EYKPIGKYLL	31.005	70.02	Unknown	Yes	SpII	No
YP_170296.1	8mer_H2_Kd.p.mtx	YIINYKLL	28.236	65.72	Unknown	No	CYT	No
YP_170390.1	10mer_H2_Kd.p.mtx	WYIYPSKIGL	33.593	75.86	Cytoplasmic membrane	No	TMH	No
YP_170399.1	9mer_H2_Kd.p.mtx	KYYKINTQL	28.272	69.21	Unknown	No	SpI	No

YP_170490.1	10mer_H2_Kd.p.mtx	KYPNAIIEGI	28.817	65.08	Cytoplasmic	No	CYT	No
YP_170502.1	8mer_H2_Kd.p.mtx	SIIFFTPL	30.095	70.05	Cytoplasmic membrane	No	CYT	No
YP_170513.1	8mer_H2_Kd.p.mtx	YYPVTAEL	29.661	69.04	Unknown	Yes	CYT	No
YP_170604.1	8mer_H2_Kd.p.mtx	YIINYKLL	28.236	65.72	Unknown	No	CYT	No
YP_170662.1	10mer_H2_Kd.p.mtx	KYPHIPIGLL	29.381	66.35	Cytoplasmic	No	CYT	No
YP_170683.1	9mer_H2_Kd.p.mtx	KYIQLFTQL	28.308	69.3	Cytoplasmic	No	CYT	No

Those highlighted in gray are good candidate epitopes. *CYT* cytoplasmic, *TMH* transmembrane helix, *SpI* signal peptide (signal peptidase I), *SpII* lipoprotein signal peptide (signal peptidase II)

Table 4
MHCII peptides identified by RankPep with binding thresholds ≤ 50 nM along with the corresponding protein localization information as identified by PSortB, SecretomeP, LipoP, and SignalP

	RankPep				PSortB	SecretomeP	LipoP	SignalP
Accession	PSSM	Consensus peptide	Optimal score	Binding threshold	Location	Secretion possible	Class	Signal peptide present
YP_003097805.1	I_Ed.p.mtx	KYFKKQALK	37.937	52.68	Cytoplasmic	No	CYT	No
YP_169178.1	I_Ed.p.mtx	KDKLRNTLK	39.793	55.26	Cytoplasmic	No	CYT	No
YP_169210.1	I_Ad.p.mtx	QMVSAAAAL	32.402	60.97	Cytoplasmic	No	CYT	No
YP_169235.1	I_Ed.p.mtx	KQLDKNTLK	36.584	50.8	Cytoplasmic	No	CYT	No
YP_169256.1	I_Ad.p.mtx	TALHLAAAE	28.526	53.68	Cytoplasmic	No	CYT	No
YP_169363.1	I_Ad.p.mtx	EAVKAAHAA	29.06	54.68	Cytoplasmic	No	CYT	No
YP_169420.1	I_Ad.p.mtx	NAVHGSDAE	27.17	51.12	Extracellular	No	CYT	No
YP_169432.1	I_Ad.p.mtx	KGVSIADPN	28.337	53.32	Cytoplasmic	No	CYT	No
YP_169500.1	I_Ad.p.mtx	TAVSHAQAG	27.347	51.46	Cytoplasmic	No	CYT	No
YP_169504.1	I_Ad.p.mtx	EMISYFQAN	26.843	50.51	Cytoplasmic	No	CYT	No
YP_169572.1	I_Ed.p.mtx	YYVKDNGKK	39.95	55.47	Unknown	No	SpI	No
YP_169652.1	I_Ad.p.mtx	EAFHATLEE	33.037	62.16	Cytoplasmic	No	CYT	No
YP_169653.1	I_Ed.p.mtx	FEQWQNRLK	38.432	53.37	Cytoplasmic	No	CYT	No
YP_169665.1	I_Ed.p.mtx	EYHKQITLE	36.367	50.5	Unknown	No	CYT	No
YP_169711.1	I_Ad.p.mtx	EIVHASFAS	26.6	50.05	Cytoplasmic	No	CYT	No
YP_169903.1	I_Ad.p.mtx	EAIRVTAAN	28.095	52.86	Cytoplasmic	No	CYT	No
YP_169914.1	I_Ed.p.mtx	KYVFLNRLH	37.497	52.07	Cytoplasmic	No	CYT	No
YP_169960.1	I_Ad.p.mtx	NQLHAAVVE	26.772	50.38	Cytoplasmic	No	CYT	No
YP_169973.1	I_Ed.p.mtx	KVKWQNHLE	38.751	53.81	Unknown	No	CYT	No

YP_169977.1	I_Ed.p.mtx	YVKKDITLK	36.744	51.02	Cytoplasmic	No	SpI	No
YP_170041.1	I_Ad.p.mtx	KAVQATQAN	33.449	62.94	Cytoplasmic	No	CYT	No
YP_170069.1	I_Ed.p.mtx	LRKWQITLK	42.771	59.39	Cytoplasmic membrane	No	TMH	No
YP_170075.1	I_Ad.p.mtx	KAQQVTHAK	28.251	53.16	Cytoplasmic	No	CYT	No
YP_170087.1	**I_Ed.p.mtx**	**HVHKQNTIK**	**41.962**	**58.27**	**Unknown**	**Yes**	**SpI**	**Yes**
YP_170105.1	I_Ad.p.mtx	RTIHVAIAN	27.163	51.11	Cytoplasmic	No	CYT	No
YP_170109.1	I_Ed.p.mtx	KLRWQPTIK	36.323	50.44	Cytoplasmic	No	CYT	No
YP_170136.1	I_Ed.p.mtx	KRVKQQRLK	47.847	66.44	Cytoplasmic	Yes	CYT	No
YP_170144.1	**I_Ed.p.mtx**	**KYAKQIAQK**	**45.31**	**62.92**	**Periplasmic**	**Yes**	**SpI**	**No**
YP_170148.1	I_Ad.p.mtx	SHVHIKHGE	27.093	50.98	Unknown	No	CYT	No
YP_170159.1	I_Ad.p.mtx	HHISIAQAN	26.595	50.04	Cytoplasmic	No	CYT	No
YP_170185.1	I_Ed.p.mtx	KEMKQNQAK	38.308	53.19	Cytoplasmic	Yes	CYT	No
YP_170193.1	I_Ad.p.mtx	CAVIATKAG	26.686	50.21	Cytoplasmic	No	CYT	No
YP_170212.1	I_Ad.p.mtx	EAVSCHDVN	26.777	50.38	Cytoplasmic	No	CYT	No
YP_170216.1	**I_Ed.p.mtx**	**IVHFQNTLK**	**46.033**	**63.92**	**Outer membrane**	**Yes**	**SpII**	**Yes**
YP_170224.1	I_Ed.p.mtx	KVKKQKTLK	44.177	61.34	Cytoplasmic	No	CYT	No
YP_170268.1	I_Ad.p.mtx	QALHVVDPE	30.635	57.64	Cytoplasmic	No	CYT	No
YP_170278.1	I_Ad.p.mtx	QWIHKALAN	33.97	63.92	Cytoplasmic	No	CYT	No
YP_170369.1	I_Ed.p.mtx	KQHKQPTQK	37.068	51.47	Cytoplasmic membrane	No	CYT	No
YP_170434.1	I_Ed.p.mtx	EYVRRQTLK	40.786	56.64	Cytoplasmic	No	CYT	No
YP_170467.1	**I_Ed.p.mtx**	**EYRKQNTQN**	**38.61**	**53.61**	**Unknown**	**Yes**	**SpI**	**Yes**
YP_170508.1	I_Ed.p.mtx	QHAKQNKLK	38.031	52.81	Cytoplasmic	No	CYT	No
YP_170509.1	**I_Ad.p.mtx**	**QGVSYLPSY**	**27.372**	**51.5**	**Outer membrane**	**Yes**	**SpII**	**Yes**
YP_170651.1	I_Ad.p.mtx	SGVHATLAG	30.03	56.51	Cytoplasmic membrane	No	TMH	No

Those highlighted in gray are good candidate epitopes

1. Generate a FASTA file with the sequences from those proteins identified in Subheading 3.2.
2. Use the websites described in Subheading 2.3, to analyze the proteins for their localization.
3. Combine these results with those of the RankPep analysis (*see* Tables 3 (MHC I) and 4 (MHC II)) in order to identify target proteins for further analysis.

3.4 IEDB Tools Analysis for MHC I and MHC II Proteins

This tool can identify further peptides within a given protein that could also be candidate MHC epitopes.

1. For general analysis, go to the website http://tools.iedb.org/main/tcell, and click either Peptide binding to MHC class I or Peptide binding to MHC class II molecules (*see* **Note 6**).
2. Fill in the webform with the required information:
   (a) A protein sequence as identified above.
   (b) Prediction method: IEDB_recommended.
   (c) Specify MHC source organism, alleles, and peptide lengths to search.
   (d) Show output: Percent below [cutoff] = 2.
3. Run analysis to output possible MHC peptides ranked by consensus result with the lower the percentile rank the better the binder (*see* **Note 7**).
4. Repeat **step 2** for each protein that was identified as a possible candidate (Table 5).

4 Notes

1. There are currently more than 30 prediction tools available for screening peptides for their ability to bind to MHC I and II molecules. These tools make use of a variety of algorithms including binding matrices, artificial neural networks, support vector machines, and partial least square function. The majority of the tools classify the peptides into binders and non-binders and also predict the binding affinity of the theoretical binders. All of these tools are described more thoroughly elsewhere [16, 17].
2. The downloaded file will contain all the genome's translated sequences in FASTA format.
3. Mouse models using BALB/c mice are the standard way to test the immune response to LVS challenges; the identified MHC I alleles for this strain of mice are H-2-Db, H-2-Dd, H-Kb, and H-Kd and for MHC II are I_Ad and I_Ed.

Table 5
Example result from IEDB MHCI analysis of YP_170216.1 (*see* Note 8)

Allele	#	Start	End	Pep Len	Sequence	Method used	% Rank	ANN IC50 (nM)	ANN Rank	SMM IC50 (nM)	SMM rank	CombLib. Sidney2008 score	ComLib. Sidney2008 rank
H-2-Kd	1	214	222	9	KYAQNQHKL	Consensus (ANN,SMM,CombLib_Sidney2008)	0.2	38.7	0.2	46.6	0.2	0.0	12.1
H-2-Kd	1	214	222	9	KYAQNQHKL	Consensus (ANN,SMM,CombLib_Sidney2008)	0.2	16.7	0.1	41.8	0.2	2e-05	0.5
H-2-Kd	1	374	382	9	FYQAYYDYI	Consensus (ANN,SMM,CombLib_Sidney2008)	0.2	38.7	0.2	46.6	0.2	0.0	12.1
H-2-Kd	1	374	382	9	FYQAYYDYI	Consensus (ANN,SMM,CombLib_Sidney2008)	0.2	16.7	0.1	41.8	0.2	2e-05	0.5
H-2-Kd	1	81	89	9	KLQVSIGNI	Consensus (ANN,SMM,CombLib_Sidney2008)	0.7	599.2	0.7	160.3	0.6	0.0	4.8
H-2-Kd	1	477	486	10	NYVDPHYDNP	SMM	0.7	0.0	0.0	2381.3	0.7	0.0	0
H-2-Ld	1	30	38	9	RPKIKAPAL	Consensus (ANN,SMM)	0.8	2190.8	1.2	133.9	0.4	0.0	0
H-2-Kd	1	188	196	9	AYFSLITAK	Consensus (ANN,SMM,CombLib_Sidney2008)	0.9	1083.1	0.9	32.7	0.2	3e-05	0.9
H-2-Ld	1	280	288	9	ANIKVPINM	Consensus (ANN,SMM)	0.9	1651.6	1.1	209.7	0.7	0.0	0
H-2-Kd	1	37	46	10	ALWNNQDKNI	SMM	1	0.0	0.0	2877.8	1.0	0.0	0

(continued)

Table 5
(continued)

Allele	#	Start	End	Pep Len	Sequence	Method used	% Rank	ANN IC50 (nM)	ANN Rank	SMM IC50 (nM)	SMM rank	CombLib. Sidney2008 score	ComLib. Sidney2008 rank
H-2-Kd	1	465	473	9	LYQALGSGY	Consensus (ANN,SMM,CombLib_Sidney2008)	1.1	1723.4	1.1	277.7	1.1	2e-05	0.6
H-2-Kd	1	377	386	10	AYYDYINTLQ	SMM	1.1	0.0	0.0	3167.3	1.1	0.0	0
H-2-Kd	1	188	197	10	AYFSLITAKE	SMM	1.2	0.0	0.0	3292.7	1.2	0.0	0
H-2-Kd	1	188	197	10	AYFSLITAKE	SMM	1.2	0.0	0.0	3247.8	1.2	0.0	0
H-2-Kd	1	214	223	10	KYAQNQHKLG	SMM	1.2	0.0	0.0	3292.7	1.2	0.0	0
H-2-Kd	1	214	223	10	KYAQNQHKLG	SMM	1.2	0.0	0.0	3247.8	1.2	0.0	0
H-2-Ld	1	151	159	9	INIARQFKL	Consensus (ANN,SMM)	1.2	1140.5	0.9	441.8	1.5	0.0	0

4. As for most of the available tools, either online or stand-alone, RankPep has a limit on the number of sequences (maximum 100 per search) it can analyze at once. Therefore, the complete FASTA file will have to be broken up into smaller chunks, or individual sequences will have to be used. The simplest way to perform this is to cut and paste the required number into a new text file.
5. The html files require a tool to parse the search results into a tab-delimited format that can be easily filtered. One such online tool can be found here: http://www.webtoolhub.com/tn561393-html-to-text-converter.aspx.
6. The website only allows one protein at a time to be analyzed. For power users, download the software from http://tools.iedb.org/main/download (MHC class I or class II predictors), and follow provided installation instructions. Each allele and each length must be searched separately, so by writing a simple script that can loop through each allele and each length available for that allele, batch analysis is possible.
7. To see the full overview of the results, check the expand results box. This shows the IC50 values for each algorithm. For baseline cutoffs, binding affinities (IC50) $\leq$ 50 nM are considered strong binders and $\leq$500 nM medium binders.
8. To download the results, click the download link at the bottom of the results page. The table is truncated due to its length.

References

1. Falk K, Rötzschke O, Stevanović S et al (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290–296
2. Chicz RM, Urban RG, Gorga JC et al (1993) Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. J Exp Med 178:27–47
3. Shen H, Harris G, Chen W et al (2010) Molecular immune responses to aerosol challenge with Francisella tularensis in mice inoculated with live vaccine candidates of varying efficacy. PLoS One 5:e13349
4. Singh H, Raghava GPS (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinformatics 19:1009–1014
5. Jurtz V, Paul S, Andreatta M, Marcatili P et al (2017) NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. J Immunol 199(9):3360–3368
6. Jensen KK, Andreatta M, Marcatili P et al (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology 154:394–406
7. Zhang L, Chen Y, Wong H-S et al (2012) TEPITOPEpan: extending TEPITOPE for peptide binding prediction covering over 700 HLA-DR molecules. PLoS One 7:e30483
8. Lata S, Bhasin M, Raghava GPS (2009) MHCBN 4.0: a database of MHC/TAP binding peptides and T-cell epitopes. BMC Res Notes 2:61
9. Reche PA, Glutting J-P, Zhang H, Reinherz EL (2004) Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. Immunogenetics 56:405–419
10. Vita R, Zarebski L, Greenbaum JA et al (2010) The immune epitope database 2.0. Nucleic Acids Res 38:D854–D862
11. Gardy JL, Spencer C, Wang K et al (2003) PSORT-B: improving protein subcellular localization prediction for Gram-negative bacteria. Nucleic Acids Res 31:3613–3617

12. Nielsen H (2017) Predicting secretory proteins with SignalP. Methods Mol Biol 1611:59–73
13. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005) Non-classical protein secretion in bacteria. BMC Microbiol 5:58
14. Bendtsen JD, Jensen LJ, Blom N et al (2004) Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17:349–356
15. Juncker AS, Willenbrock H, Von Heijne G et al (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662
16. Lin HH, Ray S, Tongchusak S et al (2008) Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. BMC Immunol 9:8
17. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S (2015) An overview of bioinformatics tools for epitope prediction: implications on vaccine development. J Biomed Inform 53:405–414



Chapter 19

Enhancing Mass Spectrometry-Based MHC-I Peptide Identification Through a Targeted Database Search Approach

Prathyusha Konda, J. Patrick Murphy, Morten Nielsen, and Shashi Gujar

Abstract

MHC-bound peptide ligands dictate the activation and specificity of CD8+ T- cells-based and thus are important for devising T-cell immunotherapies. In recent times, advances in mass spectrometry (MS) have enabled the precise identification of these peptides, wherein MS/MS spectra are compared against a reference proteome. Unfortunately, matching immunopeptide MS/MS to reference proteome databases is hindered by inflated search spaces attributed to the number of matches that need to be considered due to a lack of enzyme restriction. These large search spaces limit the efficiency with which MHC-I peptides are identified. Here we offer a solution to this problem whereby we describe a targeted database search approach and accompanying tool SpectMHC that is based on a priori predicted MHC-I peptides (Murphy et al., J Proteome Res 16:1806–1816, 2017).

Key words MHC ligandome, Mass spectrometry, Bioinformatics, Immuno-informatics, MHC peptides, NetMHC, SpectMHC

1 Introduction

Antigenic peptides, presented by major histocompatibility complexes (MHCs), play a major role in adaptive immunity and dictate the specificity of T-cell responses. These peptides are presented by either class I (to CD8+ cytotoxic T cells) or class II (to CD4+ helper T cells) MHCs [1, 2]. Class I MHCs are present on the surface of all nucleated cells, while class II MHC complexes are present on the surface of antigen-presenting cells (APCs). Class I MHC ligands are usually 8–11 amino acids long and are derived from proteins of invading pathogens, or host proteins that are considered "self," tumor-associated antigens (TAAs), or mutation-bearing tumor-specific antigens (TSAs). Class II MHC ligands are several amino acids longer than class I peptides and are typically

Prathyusha Konda and J. Patrick Murphy contributed equally to this work.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_19, © Springer Science+Business Media, LLC, part of Springer Nature 2019

derived from endocytosed proteins. Together, these peptide-bound MHC class I and class II complexes provide a "window" into the constantly changing intracellular environment and serve as a signal for surveillance by CD8+ and CD4+ T cells.

Identifying MHC peptides through mass spectrometry (MS) has enabled a more detailed understanding of immune responses [3]. In typical MHC-I peptide identification workflows, MHC-bound peptides are first immunoprecipitated using anti-MHC-I antibodies and analyzed by LC-MS/MS for their identification. Traditionally, MS/MS spectra for eluted MHC-I ligands are compared against a reference proteome database using no enzymatic digestion specificity. This lack of enzyme specificity creates large search spaces that decrease statistical power in assigning false-positive peptides and ultimately reduces the number of peptide identifications in an experiment. Here, we describe a solution to this limitation, whereby a targeted database approach is performed, aided by our database generating tool SpectMHC. The approach relies on an MHC prediction algorithm, NetMHC [4, 5], to create an MHC-I peptide-specific database thus limiting the database to the kind of peptides which are more likely to be present in the samples. We have previously shown that this targeted search strategy improves peptide identifications for both mouse and human MHC-I ligandomes by more than twofold and is superior to traditional "no enzyme" searches of reference proteome databases [6]. As such, this approach promises to uncover otherwise missed MHC-I ligands, including those containing novel T-cell epitopes of therapeutic potential.

2 Materials

2.1 System Requirements

1. A Linux or Darwin (Mac OS) system, preferably with at least 6 GB of RAM.

2.2 LC-MS/MS Data from an MHC-I Immunoprecipitation (See Note 1)

1. Raw LC-MS/MS data from immunoprecipitated MHC/HLA peptides from mouse or human tissue [2]. The targeted search approach has been validated using the mouse H-2 D^b- and H-2 K^b-specific antibodies (B22.249 and Y3 hybridomas, respectively), as well as the human pan-HLA antibody (clone W6/32) (*see* **Note 2**).

2.3 FASTA File Containing Input Reference Protein Sequences

1. Organism-specific protein reference databases can be downloaded in FASTA format from the UniProt database (www.uniprot.org) or other reputable protein sequence repository (*see* **Note 3**).
2. Optionally, a custom protein reference database can be formed from RNASeq or exome sequencing data translated into protein sequences and arranged in FASTA file format.

2.4 Locally Installed MHC/HLA Prediction Software

1. NetMHC prediction software can be freely downloaded for academic users: NetMHCpan 4.0 (http://www.cbs.dtu.dk/services/NetMHCpan/) and NetMHC 4.0 (http://www.cbs.dtu.dk/services/NetMHC-4.0/).
2. Installation instructions are available in the readme file associated with the version of NetMHC (*see* **Note 4**).

2.5 Locally Installed SpectMHC Database Compiler Tool

1. SpectMHC is free to download for academic users to compile FASTA databases from NetMHC-predicted peptides and is available at https://github.com/Prathyusha-konda/SpectMHC-v2.
2. To download the tool, open terminal at your desired location in your system and type:
   ">git clone git@github.com:Prathyusha-konda/SpectMHC-v2.git" (if using ssh authentication)
   ">git clone https://github.com/Prathyusha-konda/SpectMHC-v2.git".
3. Unzip the folder and move the files into your workspace.

2.6 Proteomics Database Search Software

1. User-preferred database search software such as Proteome Discoverer 1.4 (or greater) capable of implementing MS search algorithms and false discovery rate (FDR) correction. The targeted search approach has been validated using both Sequest and Mascot searches using Percolator to control the FDR. However, the approach is also applicable to Andromeda, MS-GF+, X!Tandem, and others.

3 Methods

3.1 Tissue MHC/HLA Typing

1. Determine the MHC or HLA haplotypes specific to the mouse or human tissue sample to be analyzed. It is preferable to perform 4-digit typing on HLA A, B, and C loci to conform with the NetMHC prediction algorithms.
2. MHC-I alleles for the common C57BL/6 mouse strain are H-2 K^b and H-2 D^b.
3. If using a different mouse background, determine the MHC haplotype using this table (http://tools.thermofisher.com/content/sfs/brochures/Mouse_Haplotype_Table.pdf).
4. MHC/HLA haplotypes for most common cell lines can be found at the TRON Cell Line Portal (http://celllines.tron-mainz.de/).
5. For individual patient samples, MHC haplotyping can be performed by PCR-based kits or from patient-specific sequencing data using a bioinformatics tool such as Seq2HLA [7], HLAreporter [8], or HLAminer [9].

3.2 Formulating an MHC/HLA-Specific FASTA Database with SpectMHC

1. Move the reference FASTA file (e.g., full human or mouse reference proteome) into the SpectMHC-v2 folder.
2. Open a terminal window from the SpectMHC folder containing your parent file on your Linux/Darwin system, and type ">bash ./spectmhc.sh -h" (this is a help command). Once executed, you will be prompted with a set of instructions about how to implement SpectMHC.
3. Implement the basic usage of SpectMHC using the command ">bash ./spectmhc.sh [-r] [-s] <NetMHC folder> <MHC version> <input fasta> <binding cutoff> <allele> <length of peptides> <number of split files>." The description of these required arguments are NetMHC folder (path to your NetMHC folder), MHC version (version of NetMHC being used), input FASTA (reference FASTA file), binding cutoff (rank cutoff to use for predicting ligands), allele (MHC/HLA allele to predict), length of peptides (if more than one length, separate them by ","), and number of split files (to specify how many split files are to be generated).
4. Set the number of split files so that 1000–2000 sequences are in each file, to reduce processing time.
5. Specify a 2% rank as a cutoff for selecting the predicted peptides. Peptides with <0.5% rank are considered strong binders, and 0.5–2% rank are weaker binders.
6. Optional arguments: Insert "-r" if you want to save raw NetMHC output files and/or "-s" if you want to split the reference FASTA file.
7. You will obtain a customized MHC-peptide targeted database in a FASTA file (*see* **Note 5**).

3.3 Performing a Database Search with a SpectMHC-Generated FASTA File

1. Load the raw LC-MS/MS data and newly generated targeted FASTA database into Proteome Discoverer or search software of choice. For duplicate peptide entries, Proteome Discoverer automatically appends all headers to a single header.
2. It is recommended to perform an initial search using no enzyme specificity against the full reference database to assess the purity of the MHC/HLA immunoprecipitation. Using the online (or offline) version of NetMHC, ensure that more than ~80% of the identified peptides from this search have <2% NetMHC rank values for the specified alleles.
3. Perform a targeted FASTA database search using "No Cleavage" enzyme specificity search options and preferred search algorithm (*see* **Notes 6** and **7**). In general, the approach works best with MS1 tolerances ≤5 ppm and MS2 tolerance ≤0.5 Da for collision-induced dissociation (CID) fragmentation (*see* **Note 8**). Perform FDR correction with Percolator or another comparable tool. In the Proteome Discoverer

Percolator node, the target-decoy approach [10] is automatically implemented, and a reversed (decoy) database does not need to be made with SpectMHC (*see* **Note 9**).

4. Report the peptides with "medium" (5% FDR) or "high" (1% FDR) confidence in the final output datasets filtered based on q-values (Percolator FDR estimates). The targeted database search should lead to significantly more peptides than the number of MHC/HLA binders (NetMHC ≤2%) in the "no enzyme" reference database search (*see* **Note 10**).
5. For quantitative analysis (label-free, SILAC, or tagging methods), ratios can be compared by grouping all spectra matching to a peptide to a single peptide group used for quantitation (automatically performed by Proteome Discoverer).

4 Notes

1. LC-MS/MS should be collected using shotgun LC-MS/MS with high-resolution (e.g., Orbitrap) MS scans (~350–700 m/z should encompass most MHC/HLA peptides) followed by data-dependent acquisition of MS/MS scans using ion trap collision-induced dissociation (CID), Orbitrap higher-energy collisional dissociation (HCD), or Orbitrap electron-transfer/higher-energy collision dissociation (EThcD). The targeted search technique has been validated with CID MS2 scans.
2. To test whether the database setup is working correctly, sample raw LC-MS/MS files from a C57BL/6 mouse MHC-I immunoprecipitate can be accessed from https://chorusproject.org/pages/index.html, ID#1098.
3. Noncanonical protein forms such as isoforms, splice variants, pseudogenes, or alternative reading frames can be included in the reference database with a minimal effect on search space compared to a "no enzyme" reference search containing all such protein variants.
4. Both NetMHC and NetMHCpan MHC/HLA binding predictions are made using artificial neural networks trained on naturally eluted and affinity binding data. NetMHC 4.0 has been trained on 81 human alleles, whereas NetMHCpan 4.0 covers 172 alleles.
5. If desired, known common contaminant peptides observed in the samples of interest (degraded intracellular or extracellular proteins) can be added to the database, so their spectra are not falsely matched to MHC/HLA peptides.
6. We find that Mascot searches with no cleavage followed by Percolator FDR correction result in a greater improvement in

the number of peptides identified compared to Sequest searches with no cleavage.

7. If implementing the "no cleavage" search using Sequest in software where this search type is not available as an option, a no cleavage search can be specified using the text "No cleavage 1 1 J" as the enzyme specificity.
8. For CID, HCD, or EThcD data collected using high-resolution Orbitrap MS/MS scans, tolerances can be decreased to less than 0.5 Da and may improve specificity.
9. Optionally, a target database can be formed by reversing the reference database and then running SpectMHC on the reversed protein database. We have shown that there is not a significant difference in peptide FDR estimation between the two target-decoy approaches.
10. In addition to the increased number of peptide identifications, MS/MS searches using targeted databases improve the search speed by ~6-fold compared to the "no enzyme" reference database searches.

Acknowledgments

We gratefully acknowledge Dr. Stefan Stevanovic, Dr. Dan Kowalewski, and Dr. Heiko Schuster (Department of Immunology, Institute for Cell Biology, University of Tubingen) for helpful discussions in devising the targeted database search approach. We also acknowledge Ms. Youra Kim for her support with editing of this document. We acknowledge financial support from the Canadian Institutes of Health Research (CIHR), Canadian Cancer Society Research Institute (CCSRI), the Beatrice Hunter Cancer Research Institute (BHCRI), and the Dalhousie Medical Research Foundation (DMRF).

References

1. Rock KL, Reits E, Neefjes J (2016) Present yourself! By MHC class I and MHC class II molecules. Trends Immunol 37(11): 724–737
2. Blum JS, Wearsch PA, Cresswell P (2013) Pathways of antigen processing. Annu Rev Immunol 31:443–473
3. Caron E, Kowalewski DJ, Koh CC et al (2015) Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. Mol Cell Proteomics 14(12):3105–3117
4. Andreatta M, Nielsen M (2015) Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32(4):511–517
5. Jurtz V, Paul S, Andreatta M et al (2017) NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. J Immunol 199(9):3360–3368
6. Murphy JP, Konda P, Kowalewski DJ et al (2017) MHC-I ligand discovery using targeted database searches of mass spectrometry

data: implications for T-cell immunotherapies. J Proteome Res 16:1806–1816

7. Boegel S, Scholtalbers J, Löwer M et al (2015) In silico HLA typing using standard RNA-Seq sequence reads. Methods Mol Biol 1310:247–258
8. Huang Y, Yang J, Ying D et al (2015) HLA reporter: a tool for HLA typing from next generation sequencing data. Genome Med 7(1):25
9. Warren RL, Choe G, Freeman DJ et al (2012) Derivation of HLA types from shotgun sequence datasets. Genome Med 4(12): 95
10. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4(3):207–214



Chapter 20

Prioritization of Therapeutic Targets of Inflammation Using Proteomics, Bioinformatics, and In Silico Cell-Cell Interactomics

Arsalan S. Haqqani and Danica B. Stanimirovic

Abstract

Protein-protein interactions play key roles in leukocyte extravasation process into the brain and have been attractive therapeutic targets for inhibiting brain inflammation using blocking (or neutralizing) antibodies. These targets include protein-protein interactions between cytokines (or chemokines) and their receptors on leukocytes and between adhesion molecules of leukocyte and brain endothelium. While a number of therapeutics against these targets are currently used in clinic for treatment of brain autoimmune and inflammatory disorders (e.g., multiple sclerosis), they are associated with side effects partly due to the off-target actions (i.e., nonspecific targets). There is a need for novel targets involved in the leukocyte extravasation process that are specific to leukocyte subsets or to individual inflammatory disorder and are amenable for drug development (i.e., druggable). We recently described the blood-brain barrier (BBB) Carta Project as a comprehensive collection of molecular "maps" consisting of multiple experimental omics (including RNA sequencing, proteomics, glycoproteomics, glycomics, metabolomics) and in silico informatics analyses on a number of mammalian species from hundreds of internal, publically available, or curated datasets. Utilizing the datasets and tools from the BBB Carta Project, we describe a methodology to identify novel "druggable" targets involving protein-protein interactions between activated leukocytes and brain endothelial cells using a combination of proteomics, bioinformatics, and in silico interactomics. The result is a prioritized list of protein-protein interactions in a network consisting of leukocyte-brain endothelial cell communication and contacts. These interactions can be further pursued for development of therapeutics such as neutralizing antibodies and their validation through preclinical testing. In addition to targeting brain inflammation, the method described here is applicable for peripheral inflammation and provides the opportunity to target important cell-cell interactions and communications that are more specific/selective for inflammatory disorders and improve currently available therapies.

Key words Protein-protein interactions, Intercellular, Target prioritization, Therapeutics, Inflammation, Extravasation, Druggable, Proteomics, Bioinformatics

1 Introduction

Proteomics and other "omics" technologies have been applied to profile genome-wide molecular changes in diseased states in the hope of identifying new targets for therapy and diagnosis. The main

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_20, © Springer Science+Business Media, LLC, part of Springer Nature 2019

bottlenecks in translating "omics" discoveries into applications are overwhelming datasets that are difficult to rationalize and prioritize for validation and development. Validating entire "lists" of differentially expressed genes/proteins is costly and time-consuming. As a result, a single-factor/reductionist approach has often been applied to select molecules for further evaluation, and a significant portion of disease-implicated molecules are simply overlooked. Prioritization of identified differentially expressed molecules in the entire dataset in an unbiased manner is therefore needed to identify novel and more specific targets.

Target prioritization involves narrowing down the candidate targets to those that can be exploited for the development of therapeutic ligand or in targeted diagnostic. Typically, the lead target is used to develop a series of targeting agents (TAs) using combinatorial small molecule libraries or monoclonal antibody generation by hybridoma or antibody display technologies; these targeting agents are then evaluated and optimized to select a "lead drug." In most cases, the target prioritization is driven by the "druggability" of the targets. Some important criteria for target "druggability" include differential expression of the target in diseased condition(s), specificity for the disease and tissue of interest, and its structural accessibility for binding to the TA for a rapid (functional) response. The later criterion is dependent on the type of TA; whereas for small molecules "druggable" targets could be intracellular, for biologic therapeutics the accessibility of targets has been generally limited to molecules secreted into the extracellular milieu or expressed on the cell surface. A number of these "accessible" molecules have been therapeutically targeted using blocking (or neutralizing) antibodies in various disorders; examples in CNS inflammation include Tysabri® (natalizumab), an anti-VLA4 blocking antibody [1], and Remicade® (infliximab), a TNFα-blocking antibody in multiple indications [2]. For both small molecule and biologic TAs, some details of the target structure and TA interactions with specific target epitopes are also desirable [3]. Due to space restriction, we have limited the method described in this chapter to prioritization of proteomics targets of inflammation to identify lead molecules that can be pursued to develop therapeutic antibodies useful in inhibiting specific steps in the inflammation process related to leukocyte recruitment in brain vasculature.

Molecules involved in the process of leukocyte extravasation have been attractive therapeutic targets for inhibiting brain inflammation. Leukocyte extravasation is a multistep process, involving the movement of leukocytes out of the circulatory system, their activation, binding, and diapedesis through vascular endothelium and their migration toward the site of tissue damage or infection [4, 5]. Protein-protein interactions play key roles in each step of the process [6–8]: leukocyte activation/chemotaxis involves cytokine interactions with leukocyte-expressed receptors

(e.g., TNFα receptor, IL23R), while leukocyte contact with and diapedesis through endothelial cells (ECs) involve cell adhesion molecules (e.g., VLA4-VCAM1, LFA1-ICAM1) [9]. While intracellular protein-protein interactions (e.g., in signal transduction) also play important roles in the process, the interactions involving "accessible" membrane and secreted molecules are considered more attractive targets for biologic therapeutics. Hence Tysabri® and Remicade® are clinically and successfully used to inhibit T cells from entering the brain by blocking VCAM1-VLA4 interaction in multiple sclerosis and to reduce leukocyte activation by preventing TNFα binding its receptor in rheumatoid arthritis, respectively. Despite their use in clinic, these therapeutics are not free of side effects [6, 10, 11], which are in part due to widespread physiological functions of their target antigens. In addition, targeting these inter actions alone does not provide complete protection, suggesting that other molecules may "compensate" when these interactions are blocked [12]. Thus, there is a need to identify new "druggable" protein-protein interactions involved in the leukocyte extravasation process that are specific to pathogenic cell type or to individual inflammatory disorder.

The project to create a comprehensive collection of molecular "maps" of the blood-brain barrier (BBB) and the neurovascular unit in efforts to identify a repertoire of "druggable" molecules that can be used to selectively target brain disorders, also known as BBB Carta Project (or the Neurovascular Unit Carta Project), was recently described [13]. These molecular maps are composed of multiple experimental omics (RNA-seq, proteomics, glycoproteomics, glycomics, metabolomics) and in silico informatics analyses. The datasets are experimental analyses on various cell types, tissues, and body fluids from human, mouse, rat, nonhuman primates, and dogs and comprise of >120 National Research Council (NRC) datasets and hundreds of publically available and curated datasets.

One of the aims of the BBB Carta Project is to identify protein-protein interactions among different constituents of the neurovascular unit [13] as they may contain novel set of "druggable" targets for treatment of various brain disorders. The protein-protein interactions might include direct (contacts) or indirect (paracrine) interactions of ECs with astrocytes, leukocytes (Th17), neurons, pericytes, or other cell types in the brain. We have proposed tools to identify novel set of "druggable" protein-protein interactions between Th17 and ECs using a combination of proteomics, bioinformatics, and in silico interactomics [8, 14] (Fig. 1). Here we provide updated protocols for this method. The method involves inflammatory activation of brain EC and T cells followed by advanced label-free proteomics (nanoLC-MS/MS) of the cells. An initial list of accessible proteins (i.e., present on cell surface and secretome) is generated and prioritized based on bioinformatics

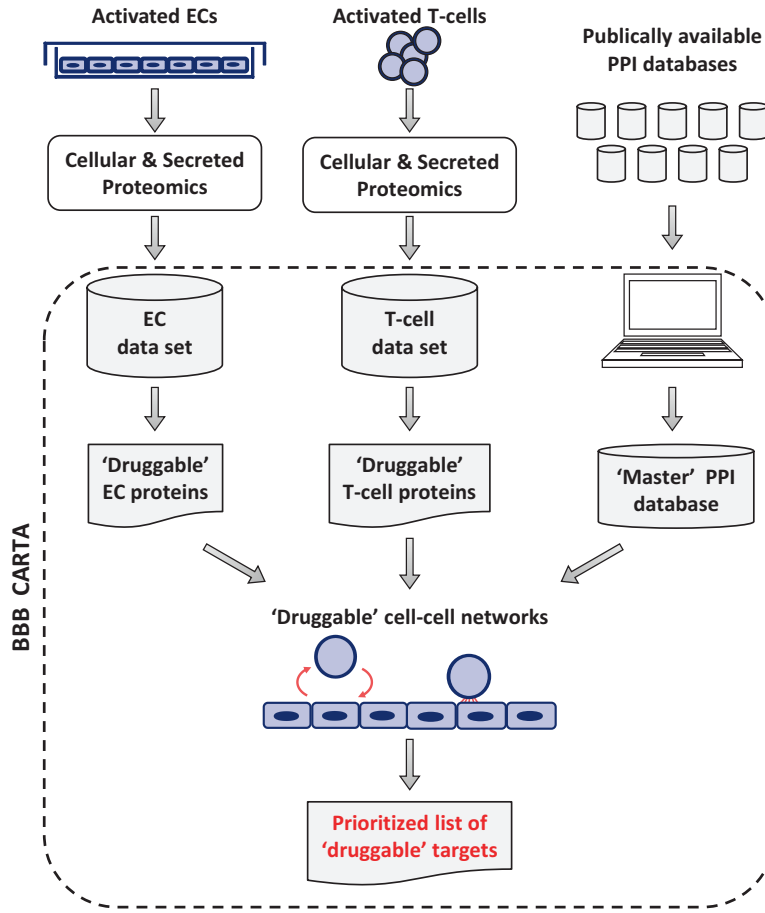


Fig. 1 Schematic representation of the experimental approach. Endothelial cells (ECs) and T cells are activated using appropriate inflammatory conditions, and their cellular and secreted proteins are analyzed by proteomics to generate a list of respective datasets. "Druggable" proteins in each dataset are identified using bioinformatics tools from BBB Carta described in the Methods. In parallel, protein-protein interaction (PPI) datasets are obtained from publicly available sites and combined to create a "master" PPI database in house. These are used to create "druggable" cell-cell networks between EC and T cells and produce a list of prioritized "druggable" targets

analysis of their potential interactions from the database of protein-protein interactions, compiled from existing databases from the public domain. The result is a reconstructed brain EC-T-cell network (interactome) consisting of cell-cell communications (paracrine) and cell-cell contacts. The network is then used to identify novel and specific protein-protein interactions that can be potentially blocked to disrupt the extravasation process. While the method has been used to identify interactions between brain ECs and T cells, it is also applicable to other disease-specific leukocyte-EC interactions as well as other cell-cell interactions and communications.

2 Materials

2.1 ECs and T Cells

1. Brain ECs used were the hCMEC/D3 human brain endothelial cell line as a stable human in vitro model of the BBB [15]. hCMEC/D3 were activated under serum-free conditions using various inflammatory insults, including TNFα/INFγ, IL-1β, or simulated ischemia/reperfusion conditions as previously described [12, 16, 17].
2. T cells with encephalo-tropism were generated by activation of lymphocytes, isolated from multiple sclerosis patients, with IL-23 as previously described [18].

2.2 Proteomic Sample Preparation

1. At least 100 μg of protein is needed from ECs and T cells for proteomics analysis.
2. Buffer exchange/concentrating columns: such as Centriprep or Amicon Ultra columns with a nominal molecular cutoff of 5000 Da.
3. Prechill acetone by refrigerating at −20 °C for 20 min.
4. Benchtop centrifuge: for example, Eppendorf Centrifuge 5415D.
5. AMBIC solution: 50 mM ammonium bicarbonate.
6. 550 Sonic Dismembrator.
7. Denaturing sodium dodecyl sulfate (SDS) buffer (DS buffer): 50 mM Tris–HCl, pH 8.5, 0.1% SDS.
8. DTT stock: 100 mM of dithiothreitol freshly prepared in 50 mM AMBIC solution.
9. IAA stock: 250 mM of iodoacetamide freshly prepared in 50 mM AMBIC solution.
10. Milli-Q water: Milli-Q® Ultrapure MS-grade water.
11. Trypsin Gold, Mass Spectrometry Grade.
12. Pierce Strong Cation Exchange Spin Columns.
13. Phosphate Buffered Saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4 \cdot 2H_2O$, 2.0 mM KH_2PO_4, pH 7.4.
14. 0.22 μm syringe filters.
15. High-pressure liquid chromatography (HPLC) grade acetonitrile.
16. HPLC grade formic acid.

2.3 Mass Spectrometry

1. An online nanoflow liquid chromatography (nanoLC) system with a reverse phase nanoLC column such as the highly reproducible nanoAcquity UPLC system (Waters) with a 100 μm inner diameter (I.D.) × 10 cm, 1.7 μm BEH130 C18 column (Waters).
2. A tandem MS instrument capable of performing electrospray ionization directly on the eluting peptides followed by high-

throughput MS/MS analysis such as the LTQ XL™, the LTQ Orbitrap™ (Thermo), or the Q-TOF™ Ultima (Waters). The LTQ instruments are of choice due to its ability to do high-throughput sequencing.

2.4 Software and Databases for Data Analysis

1. NanoLC-MS/MS data may be converted into a number of data formats that are compatible for analysis by software such as MSight, Mascot, or MatchRx [21]. The software that comes with the MS instrument or a number of available tools from http://tools.proteomecenter.org can also be used. ReAdW from http://tools.proteomecenter.org was used here for LTQ data.
   (a) MSight software is a visualization tool [19] available free of charge from https://web.expasy.org/MSight. It allows graphical representation of the LC-MS and LC-MS/MS data. MSight version 2.0 was used here.
   (b) Mascot® software (Matrix Science Ltd.) is a probability-based search engine for identifying peptide sequences from the nanoLC-MS/MS data using protein database searching [20]. Mascot® version 2.2.0 was used here.
   (c) MatchRx software (from NRC) extracts peptide abundance values from LC-MS and LC-MS/MS data and allows quantitative comparison of peptide levels in two or more samples [21]. MatchRx also overlays the quantitative differences of peptides and LC-MS/MS identification results (from Mascot) into MSight images. MatchRx version 4.0 was used in studies described here.
2. MeV or similar software for analysis, visualization, and data mining of large-scale omic data from http://mev.tm4.org.
3. SwissProt protein database in FASTA format (uniprot_sprot.fasta) and UniProt Knowledgebase (uniprot_sprot.dat) from ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete.
4. Protein-protein interaction software listed in Table 1.
5. Cytoscape an open-source software for complex network analysis and visualization [22] from http://www.cytoscape.org.
6. Although optional, access to BBB Carta Database (NRC-Human Health Therapeutics (HHT) Research Centre, Canada).

3 Methods

3.1 Proteomic Analyses of Activated ECs and T Cells

The first step requires generation of a list of proteins that are expressed in activated ECs and activated leukocytes. It is important that the right types of cells (e.g., brain EC, lung EC, Th17, B-cells) and cell-activation paradigms are chosen for each specific

Table 1
Example of a list of publicly available mammalian protein-protein interaction databases

Database	Website
BioGRID	http://thebiogrid.org
DIP	http://dip.doe-mbi.ucla.edu/dip/Main.cgi
GeneMania	https://genemania.org/
GeneRif	ftp://ftp.ncbi.nlm.nih.gov/gene/GeneRIF
HPID	http://wilab.inha.ac.kr/hpid
HPRD	http://www.hprd.org
HUGE	http://www.kazusa.or.jp/huge
IntAct	http://www.ebi.ac.uk/intact/index.html
KEGG	http://www.genome.jp/kegg
MINT	http://mint.bio.uniroma2.it
MPPI	http://mips.helmholtz-muenchen.de/proj/ppi
Reactome	http://www.reactome.org
STRING	https://string-db.org

inflammatory condition. Advanced proteomics methods can be used to identify novel proteins that are expressed specifically in the type of ECs or leukocytes being examined (*see* **Note 1**). In this section, we provide a general protocol for isolation of proteins from cellular and secreted fractions followed by proteomics analysis (*see* **Note 2**). Four types of samples are recommended for use in subsequent interactomics analysis: (1) cellular fractions of activated ECs; (2) cellular fractions of activated T cells; (3) secreted fraction of activated ECs; and (4) secreted fraction of activated T cells. For each sample type, the use of an appropriate control is highly recommended (e.g., untreated control, unrelated ECs or T cells) to obtain a more specific target.

3.1.1 Cell Activation and Harvest

1. Grow at least 1×10^6 cells on culture plates or flasks, and activate the cells to mimic inflammatory conditions, using appropriate controls (*see* **Note 3**). Examples of EC activation include inflammatory cytokines (IL1beta or TNFalpa/IFNgamma) or simulated in vitro ischemia/reperfusion as previously described [12, 16, 17]. An example of activated leukocyte includes generation of encephalotropic T cells by activation of lymphocytes (from multiple sclerosis patients) with IL-23 to produce IL-17-producing T-helper (Th17) [18].

2. Remove the medium for isolation of secreted proteins, and harvest the plated cells by scraping in PBS for isolation of cellular proteins (*see* **Note 4**).

3.1.2 Protein Isolation from Cultured Media

1. Filter the collected cultured medium through a 0.22 μm syringe filter to remove any floating cells.
2. Concentrate and desalt the medium to about less than 0.5 mL using buffer exchange/concentrating columns using the manufacturer's protocol.
3. Precipitate the proteins by adding ten volumes of ice-chilled acetone and incubating at −20 °C for at least 1 h.
4. Centrifuge at 1000 × *g* for 10 min to pellet the proteins.
5. Decant the supernatant.
6. Resuspend the pelleted secreted proteins in AMBIC solution and measure protein concentration.

3.1.3 Protein Isolation from Cells

1. Centrifuge the harvested cells at 500 × *g* for 5 min and aspirate off the PBS supernatant.
2. Resuspend the pellet in 0.1 mL of DS buffer.
3. Lyse by sonication for 4–5 cycles of 15 s each at a low setting of 3, followed by 20 s on ice.
4. Centrifuge at 10,000 × *g* for 10 min at 4 °C to pellet the cell debris and nuclei.
5. Transfer supernatant to a fresh tube, and measure protein concentration (e.g., by using a commercially available standard protein concentration kit).

3.1.4 Protein Digestion, Fractionation, and NanoLC-MS/MS Analysis

1. To cellular and secreted proteins, add freshly prepared DTT stock to a final concentration of 4 mM, and incubate at 95 °C for 10 min to reduce the disulfide bonds.
2. Cool at room temperature for 2 min.
3. Add freshly prepared IAA stock to a final concentration of 10 mM, and incubate at room temperature for 20 min in the dark to alkylate cysteine residues.
4. Add trypsin (in Milli-Q® water) at a 1:50 (w/w) trypsin-to-protein ratio, and incubate at 37 °C for 12–16 h.
5. Fractionate the tryptic peptides using strong cation exchange columns using the manufacturer's protocol.
6. Samples may be stored at −80 °C for up to 4 weeks or analyzed directly by LC-MS/MS.
7. Inject 1–5% of each sample into a nanoLC system setup online to a tandem mass spectrometer. For example, a typical analysis would involve peptide separation on the nanoAcquity system by gradient elution (1–95% ACN, 0.2% formic acid) over

Acc	GN	Protein name	No of peptides	Maximum Mascot score	Normalized Intensities								Median Fold-Change: $Log_2(T/C)$	Rankings					Final Ranking
					Treated$_1$	Treated$_2$	Treated$_3$	...	Control$_1$	Control$_2$	Control$_3$	...		Inflammatory response	"Druggability"	Interactomics			
																Degree	Betweeness	...	
Acc_1	GN_1	$Protein_1$	NP_1	Sc_1	#	#	#	...	#	#	#	...	MFC_1	#	#	#	#	...	#
Acc_2	GN_2	$Protein_2$	NP_2	Sc_2	#	#	#	...	#	#	#	...	MFC_2	#	#	#	#	...	#
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
P19320	VCAM1	Vascular cell adhesion protein 1	60	122	#	#	#	...	#	#	#	...	18.7	1	1	2	5	...	2
P05362	ICAM1	Intercellular adhesion molecule 1	50	105	#	#	#	...	#	#	#	...	7.1	2	1	4	7	...	8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Acc_N	GN_N	$Protein_N$	NP_N	Sc_N	#	#	#	...	#	#	#	...	MFC_N	#	#	#	#	...	#

Fig. 2 An example of a quantitative protein list showing various types of rankings

60 min at a flow rate of 400 nL/min and acquiring MS/MS spectra on 2+, 3+, and 4+ charged precursor ions with m/z values between 400 and 2000 (*see* **Note 5**).

8. Search the acquired MS/MS spectra against a human SwissProt protein database using Mascot® or another search engine (*see* **Notes 6** and **7**).
9. Obtain normalized relative quantitation on peptide and protein levels in each sample using MatchRx or another label-free quantitative software (*see* **Note 8**).
10. Generate a quantitative protein list for each sample type: (1) cellular fractions of activated ECs; (2) cellular fractions of activated T cells; (3) secreted fraction of activated ECs; and (4) secreted fraction of activated T cells. An example outline of the protein list is shown in Fig. 2.
11. Optionally, the quantitative lists of proteins are imported into the BBB Carta Databases to allow access to additional tools and datasets to identify list of "druggable" proteins.

3.2 List of "Druggable" Proteins in Activated ECs and T Cells

The purpose of this step is to prioritize the protein list by giving higher ranking to proteins that are more easily accessible to therapy (i.e., proteins on cell surface and extracellular space) and are induced under activated cell state (i.e., inflammation-responsive proteins) (*see* **Note 9**).

3.2.1 Cell Surface and Extracellular Proteins

1. Download an updated UniProt KB database, or directly query www.uniprot.org (*see* **Note 6**).
2. Use the quantitative protein list for each sample type as an initial list (Fig. 2) (*see* **Note 10**).
3. UniProt KB is used to identify if each protein is known to be present on cell surface and/or extracellular environment (*see* **Note 11**). To do this, query the protein accession in the database, and look for the following attributes:

(a) Under section "Comments" or "CC" and subsection "Subcellular Location," determine whether the protein is located in plasma membrane or is secreted.

(b) Under section "Gene Ontology" or "DR GO" and subsection "Cellular Component" or "C," determine whether the protein is located in plasma membrane or extracellular space.

(c) Under section "Sequence Annotation (Features)" or "FT" and subsection "Regions," determine whether the "Topological domain" is "Extracellular" and protein contains "Transmembrane" domain.

4. Add a new column for "ranking by druggablility" to each protein list and rank as follows (*see* **Note 12**):

   (a) Proteins with known extracellular domains found on the plasma membrane or secreted proteins are given a ranking of 1.

   (b) Other plasma membrane proteins are given a ranking of 2.

   (c) Protein with no known subcellular localizations are labeled as "Unknowns" and given a ranking of 5.

   (d) The remaining proteins (including nuclear and mitochondrial proteins) are either given a ranking of 1000 or removed from the list.

3.2.2 Inflammation-Responsive Proteins

1. For each sample type, calculate a median fold-change value for each protein in the list (*see* **Note 13**).
2. Carry out statistical analysis (e.g., t- or U-tests) to identify whether the "control" and "treated" intensities are significantly different ($p < 0.05$) (*see* **Note 14**).
3. Depending on the spread of the distribution, usually proteins showing fold change ≥ 2 and p-value ≤ 0.05 can be considered inflammation-responsive.
4. Add a new column for "ranking by inflammatory response" to each protein list, and rank as follows (*see* **Note 12**):

   (a) Proteins showing tenfold change are given a ranking of 1.

   (b) Proteins showing fourfold change are given a ranking of 2.

   (c) Proteins showing twofold change are given a ranking of 3.

   (d) Proteins showing no change are given a ranking of 1000.

3.3 Prioritization of "Druggable" Targets Using Cell-Cell Interactomics

In this step, the list of "druggable" inflammatory proteins is further prioritized based on their potential involvement in interaction between ECs and T cells. This requires the use of a database of known protein-protein interactions, which is first compiled from existing databases in the public domain. The database is used to reconstruct networks between ECs and T cells consisting of cell-cell communications and cell-cell contacts (*see* **Note 9**).

3.3.1 Protein-Protein Interaction Databases

1. A number of publicly available protein-protein interaction databases currently exist. Download each protein-protein interaction database listed in Table 1 to a local location (*see* **Note 15**).
2. Since each database has a different format, convert each database into a single format as shown in Fig. 3 (*see* **Note 16**).
3. Merge all the databases into one "master" database (Fig. 3).
4. It may be necessary to limit the interactions to immunoprecipitation and affinity pull-down assays and to mammalian systems to reduce the incidence of false interactions.
5. The resulting protein-protein interaction database, referred as "master" PPI database, is used in the next section.

3.3.2 In Silico Reconstruction of EC-T-Cell Interactions

1. Use the list of "druggable" cell surface proteins from activated ECs as "List 1" and the list of "druggable" cell surface proteins from activated T cells as "List 2."
2. Search each protein from List 1 in the "Protein" column of the "master" PPI database.
3. If the protein is found in the "Protein" column and has protein(s) in the interactant column, search each of the interactants in List 2 proteins.
4. For each interactant found, put each protein-protein interaction pair in a tab-delimited SIF file format for Cytoscape (e.g., VCAM1<tab>pp.<tab>VLA4 < newline>) (*see* **Note 17**).
5. Visualize all the protein-protein interaction pairs using Cytoscape.
6. Graph theories could be applied using Cytoscape tools or other methods [22, 23] to rank the protein list based on various types of measures of centrality (e.g., betweenness, degree) of the interactions in the network. Thus, for each

Proteins	Interactants	Pubmed Reference [optional]	Evidence [optional]	known/predicted [optional]	Database [optional]				
					HPRD	GeneMania	SRING	GeneRif	BioGrid
VCAM1									
	VLA4 (ITGA4/ITGB1)	12097820,19057841		known	✓	✓	✓	✓	✓
	ITGAD	10438935		known	✓	✓	✓	✓	✓
	ITGB7	12682249		known	✓	✓	✓	✓	
	CCL17	12193695	Reconstituted Complex	known	✓	✓		✓	✓
	CCL22			known	✓			✓	✓
	EZR	12082081	Affinity Capture-Western; Reconstituted Complex	known	✓	✓	✓	✓	✓
	MSN			known	✓	✓	✓	✓	✓
	CTSG			known	✓			✓	
	ELANE			known	✓			✓	
	IL13			known	✓			✓	
	IL1B1			known	✓				
	SEC61B			known				✓	
ICAM1									
	LFA1 (ITGAL/ITGB2)	10493852,11786177		known	✓	✓	✓	✓	✓
	Mac1 (ITGAM/ITGB2)	9488691,11073102							
	EGER								

Fig. 3 A simple format for protein-protein interaction database

measure, a new column is added to the list, and proteins are ranked by descending measure. Proteins not involved in protein-protein interactions are either removed or given a ranking of 1000 (*see* **Note 12**).

7. Optionally all the different columns of rankings can be combined to obtain a final ranked list (e.g., using reciprocal or other types of ranking methods [24–26]) (*see* **Note 18**).
8. The above steps will identify interacting proteins between ECs and T cells involved in cell-cell contacts. To identify cytokine-receptor interactions involved in cell-cell communications, the steps can be repeated using secreted proteins as one of the lists. That is, use the list of "druggable" cell surface proteins from activated ECs as "List 1" and the list of secreted proteins from activated T cells as "List 2," and vice versa, i.e., cell surface T-cell proteins as "List 1" and secreted EC proteins as "List 2."

3.4 Validation Studies

A range of protocols, from nanoLC-MS/MS to cell-cell interactomics, have been described here, and it is recommended to include various types of validation studies to have confidence in the results.

1. Quantitative results from nanoLC-MS/MS analysis are generated in an automated manor using label-free quantitative software. While validating all results is tedious, it is recommend that proteins showing statistically significant expression changes in their intensities (inflammation-responsive proteins) should be either manually validated in the raw nanoLC-MS/MS data or "visually" validated using the MSight software that allows reconstruction of the raw data into a 2D gel-like image.
2. Since nanoLC-MS/MS proteomics can produce a large list of inflammation-responsive proteins, validations of a few of the results using alternative methods are highly recommended. For confirmation of changes at protein level, enzyme-linked immunosorbent assay (ELISA), western blots, multiple reaction monitoring (MRM), or in situ immunochemistry may be used. For confirmation of changes at mRNA levels, quantitative reverse transcription polymerase chain reaction (RT-PCR), northern blots, or in situ hybridization may be used. Additionally, confirmation through literature mining also adds key value to the results by providing further validity to the method. In brain ECs we have validated a large number of proteins responding significantly to inflammatory insults by western blotting and immunochemistry, including ICAM1, VCAM1, integrins [8, 13], and by MRM (unpublished data).
3. It may be necessary to experimentally validate the found protein-protein interaction pairs between EC and T cells using molecular techniques in co-culture experiments. This requires establishing

co-cultures by culturing either EC with T cells or EC with T-cell-conditioned media, whereas EC without T cells or the conditioned media can be used as control, respectively. The co-cultures are then immunoprecipitated with the first protein, and the second protein is examined for using western blotting or MRM. Alternatively, immunocytochemistry for both proteins can be carried out to demonstrate co-localization [12].

4. Finally, to validate that the found protein-protein interaction pairs are functionally important for the extravasation process, either interaction-blocking therapeutic or knockout animal models could be used. This could include a commercial or a newly developed antibody, a peptide, or small molecule and could be tested for its ability to inhibit leukocyte chemotaxis, adhesion, and transmigration through ECs in in vitro BBB assays and various animal models [12, 27–31].

3.5 Conclusions and Further Steps

Proteomics and other genomics methods generate an overwhelming number of molecules, and prioritizing them to identify those implicated in disease is a challenge. While bioinformatics and statistical analyses can reduce the number of proteins involved, additional methods to identify novel and more specific targets are needed. We have described here a methodology to further narrow down proteomics list of activated EC and activated T cells to identify potential therapeutic targets of inflammation using in silico interactomics. The methodology is a component of the BBB Carta Project to identify protein-protein networks among different constituents of the neurovascular unit (e.g., EC, astrocytes, T cells) as novel "druggable" targets. These methods have been applied previously to reconstruct intercellular interactions and communications between activated human brain EC and activated human Th17 cells [8], identifying a large number of known and novel "druggable" protein-protein interaction pairs. Some of these include VCAM1-VLA4, ICAM1-LFA1, ICAM1-Mac1, PECAM1-PECAM1, and E-selectin-PSGL1. These interactions could be prevented or suppressed using specific antibodies against one or more interacting proteins. These antibodies could be screened in EC-T-cell co-culture experiments to demonstrate inhibition of interactions as described [12, 27–31]. This method identified several known interactions, including VCAM1-VLA4, already targeted by clinically used therapeutic antibodies (e.g., Natalizumab—Tysabri), providing a proof-of-concept for potential clinical translation of the proposed approach. In addition, it has identified novel interactions (unpublished data) providing the opportunity to identify and target other important leukocyte-EC interactions that are more specific/selective for inflammatory cell subtypes and for various types of inflammatory diseases.

4 Notes

1. Alternatively, a protein list from published proteomics analyses of activated ECs and/or activated leukocytes can also be used. In such case, the proteomics step can be skipped. However, the type of cells, activation protocols, and controls used in the publication should be noted to ensure it is appropriate for the inflammatory condition in question.
2. Subcellular proteomics can also be used to identify more specific and relevant targets such as luminal membrane proteins (blood-accessible targets), glycoproteins, lipid raft proteins, or endosomal proteins [12, 32, 33]. This would require subcellular fractionation and glycoprotein enrichment followed by the proteomics method described here.
3. Once cells have been plated and reached confluence, it is important that the cells be incubated for an appropriate period of time to allow enough proteins to accumulate in the medium for detection by proteomics. It may be necessary to measure protein content at the end of the incubation to ensure that there is enough protein (>150 μg). Usually one 10 mm plate of confluent cells gives about 100–500 μg of protein, and multiple plates may be pooled if necessary. Cells should be activated in a media that is free of any serum or other supplements. To ensure this, cells should be washed at least three times with a buffered saline solution and then incubated in serum-free medium in the presence of activating inflammatory cytokines.
4. Cell harvesting by scraping is recommended rather than by trypsin treatment to avoid digestion and "loss" of cell surface molecules.
5. Injecting at least one blank run is recommended in between samples to clean out the column and prevent cross-contamination between samples. If the file size is large, it may be necessary to split the files into two m/z ranges 400–1200 and 1200–2000.
6. An updated SwissProt database in FASTA format (uniprot_sprot.fasta) and its corresponding UniProt knowledgebase (uniprot_sprot.dat) should be downloaded monthly from UniProt FTP site (ftp://ftp.uniprot.org). The MS/MS data should be searched against the most up-to-date FASTA file.
7. For Mascot analysis, specify trypsin enzymatic cleavage with one possible missed cleavage; allow variable modification of oxidation (+15.99 Da) at the Met residues, allow fixed modification of carbamidomethyl (iodoacetamide derivative; +57.02 Da) at the Cys residues, and set parent ion tolerance ≤0.5 Da and fragment ion tolerance ≤0.2 Da.

8. Automated steps in MatchRx software allow quantitative comparison of peptide and protein levels in multiple samples [21]. This involves peak identification and quantitation (area under the curve) in each LC-MS run, alignment of multiple LC-MS runs, peak normalization and identification of differentially expressed peptides, and overlaying of the results on MSight images [19] for visual verification.
9. While these analyses can be done manually for a short protein list (<10), use of a computational method (e.g., macros, Perl, Matlab scripts) is highly recommended for larger datasets.
10. While the protein list can be from the nanoLC-MS/MS analysis in Subheading 3.1, pre-existing proteomics lists from published datasets, public and private databases can also be used. For example, various existing proteomics datasets on different neurovascular cell types in NRC-HHT BBB Carta Database can also be used as starting protein lists.
11. UniProt KB is a curated knowledgebase that contains experimental and predicted information (from species orthologs) on extracellularly exposed proteins. However, other datasets containing additional experimental evidence if available can also be used, e.g., experimental datasets on luminal proteins of ECs [13] and of in vivo vessels (unpublished) in BBB Carta Database.
12. The rankings given here are only suggestive. Other scores should be tested to identify the most optimal ranking system for each paradigm.
13. For each control-treated pair, a fold-change ratio can be calculated as base-2 log of IntensityTreated/IntensityControl and can be used to calculate a median fold-change ratio. Alternatively, median intensities can be calculated for each treated and control samples, and then a fold-change ratio value can be calculated using these values.
14. This can be done using MeV (http://www.tm4.org/mev) or similar software. If comparing two states, a t-test can be used for parametric data or a Mann-Whitney U-test for nonparametric data. If more than two states are being compared, ANOVA followed by Bonferroni post hoc test should be used.
15. It is recommended that like protein databases, the interaction databases should be routinely updated since new interactions are continuously identified and validated as well as some old ones are occasionally refuted and deleted from the databases.
16. The format can be simple or complex, although it should include two main columns: “Proteins” and “Interactant.”
17. While SIF is a simpler format and is useful for a quick “look” at the interactions using Cytoscape, use of XML formats is

recommended since additional information such as subcellular localization, expression, rankings can be included. The later data type also gives more flexibility in Cytoscape to "design" a cell-cell interaction [22].

18. We recommend testing a number of sorting/filtering methods to generate multiple lists of prioritized targets which are then compared to further narrow down the list. Examples of some of these methods (currently also in the BBB Carta Database) include weighted scoring, sum of rankings, k-mean clustering, and pattern matching.

References

1. Singer BA (2017) The role of natalizumab in the treatment of multiple sclerosis: benefits and risks. Ther Adv Neurol Disord 10:327–336
2. Benucci M, Saviola G, Manfredi M, Sarzi-Puttini P, Atzeni F (2012) Tumor necrosis factors blocking agents: analogies and differences. Acta Biomed 83:72–80
3. Pérot S, Sperandio O, Miteva MA, Camproux A-C, Villoutreix BO (2010) Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. Drug Discov Today 15:656–667
4. Rossi B, Constantin G (2008) Anti-selectin therapy for the treatment of inflammatory diseases. Inflamm Allergy Drug Targets 7:85–93
5. Engelhardt B, Wolburg H (2004) Mini-review: transendothelial migration of leukocytes: through the front door or around the side of the house? Eur J Immunol 34:2955–2963
6. Coisne C, Lyck R, Engelhardt B (2007) Therapeutic targeting of leukocyte trafficking across the blood-brain barrier. Inflamm Allergy Drug Targets 6:210–222
7. Wittchen ES (2009) Endothelial signaling in paracellular and transcellular leukocyte transmigration. Front Biosci (Landmark Ed) 14:2522–2545
8. Haqqani AS, Stanimirovic DB (2011) Intercellular interactomics of human brain endothelial cells and th17 lymphocytes: a novel strategy for identifying therapeutic targets of CNS inflammation. Cardiovasc Psychiatry Neurol 2011:175364
9. Filippi M-D (2016) Mechanism of diapedesis. Adv Immunol 129:25–53
10. Ransohoff RM (2005) Natalizumab and PML. Nat Neurosci 8:1275–1275
11. Scheinfeld N (2004) A comprehensive review and evaluation of the side effects of the tumor necrosis factor alpha blockers etanercept, infliximab and adalimumab. J Dermatolog Treat 15:280–294
12. Cayrol R, Wosik K, Berard JL, Dodelet-Devillers A, Ifergan I, Kebir H, Haqqani AS, Kreymborg K, Krug S, Moumdjian R, Bouthillier A, Becher B, Arbour N, David S, Stanimirovic D, Prat A (2008) Activated leukocyte cell adhesion molecule promotes leukocyte trafficking into the central nervous system. Nat Immunol 9:137–145
13. Haqqani AS, Stanimirovic DB (2017) Proteomes and biomarkers of the neurovascular unit. In: Caplan LR, Biller J, Leary MC, Lo EH, Thomas AJ, Yenari M, Zhang JH (eds) Primer on cerebrovascular diseases. Academic, San Diego, pp 346–350
14. Haqqani AS, Stanimirovic DB (2013) Prioritization of therapeutic targets of inflammation using proteomics, bioinformatics, and in silico cell-cell interactomics. Methods Mol Biol 1061:345–360
15. Weksler BB, Subileau EA, Perrière N, Charneau P, Holloway K, Leveque M, Tricoire-Leignel H, Nicotra A, Bourdoulous S, Turowski P, Male DK, Roux F, Greenwood J, Romero IA, Couraud PO (2005) Blood-brain barrier-specific properties of a human adult brain endothelial cell line. FASEB J 19: 1872–1874
16. Stanimirovic D, Shapiro A, Wong J, Hutchison J, Durkin J (1997) The induction of ICAM-1 in human cerebromicrovascular endothelial cells (HCEC) by ischemia-like conditions promotes enhanced neutrophil/HCEC adhesion. J Neuroimmunol 76:193–205
17. Haqqani AS, Kelly J, Baumann E, Haseloff RF, Blasig IE, Stanimirovic DB (2007) Protein markers of ischemic insult in brain endothelial cells identified using 2D gel electrophoresis and ICAT-based quantitative proteomics. J Proteome Res 6:226–239

18. Kebir H, Kreymborg K, Ifergan I, Dodelet-Devillers A, Cayrol R, Bernard M, Giuliani F, Arbour N, Becher B, Prat A (2007) Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. Nat Med 13:1173–1175
19. Palagi PM, Walther D, Quadroni M, Catherinet S, Burgess J, Zimmermann-Ivol CG, Sanchez J-C, Binz P-A, Hochstrasser DF, Appel RD (2005) MSight: an image analysis software for liquid chromatography-mass spectrometry. Proteomics 5:2381–2384
20. Hirosawa M, Hoshida M, Ishikawa M, Toya T (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. Comput Appl Biosci 9:161–167
21. Haqqani AS, Kelly JF, Stanimirovic DB (2008) Quantitative protein profiling by mass spectrometry using label-free proteomics. Methods Mol Biol 439:241–256
22. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, Pico AR, Bader GD, Ideker T (2012) A travel guide to Cytoscape plugins. Nat Methods 9:1069–1076
23. Barabási A-L, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. Nat Rev Genet 5:101–113
24. Aßfalg A, Erdfelder E (2012) CAML—maximum likelihood consensus analysis. Behav Res Methods 44:189–201
25. Abi-Haidar A, Kaur J, Maguitman A, Radivojac P, Rechtsteiner A, Verspoor K, Wang Z, Rocha LM (2008) Uncovering protein interaction in abstracts and text using a novel linear model and word proximity networks. Genome Biol 9:S11
26. Melnik O, Vardi Y, Cun-Hui Zhang C-H (2004) Mixed group ranks: preference and confidence in classifier combination. IEEE Trans Pattern Anal Mach Intell 26:973–981
27. Larochelle C, Uphaus T, Broux B, Gowing E, Paterka M, Michel L, Dudvarski Stankovic N, Bicker F, Lemaître F, Prat A, Schmidt MHH, Zipp F (2018) EGFL7 reduces CNS inflammation in mouse. Nat Commun 9:819
28. Larochelle C, Cayrol R, Kebir H, Alvarez JI, Lécuyer M-A, Ifergan I, Viel É, Bourbonnière L, Beauseigle D, Terouz S, Hachehouche L, Gendron S, Poirier J, Jobin C, Duquette P, Flanagan K, Yednock T, Arbour N, Prat A (2012) Melanoma cell adhesion molecule identifies encephalitogenic T lymphocytes and promotes their recruitment to the central nervous system. Brain 135:2906–2924
29. Ifergan I, Kebir H, Terouz S, Alvarez JI, Lécuyer M-A, Gendron S, Bourbonnière L, Dunay IR, Bouthillier A, Moumdjian R, Fontana A, Haqqani A, Klopstein A, Prinz M, Lõpez-Vales R, Birchler T, Prat A (2011) Role of ninjurin-1 in the migration of myeloid cells to central nervous system inflammatory lesions. Ann Neurol 70:751–763
30. Schulze-Topphoff U, Prat A, Prozorovski T, Siffrin V, Paterka M, Herz J, Bendix I, Ifergan I, Schadock I, Mori MA, Van Horssen J, Schröter F, Smorodchenko A, Han MH, Bader M, Steinman L, Aktas O, Zipp F (2009) Activation of kinin receptor B1 limits encephalitogenic T lymphocyte recruitment to the central nervous system. Nat Med 15:788–793
31. Podjaski C, Alvarez JI, Bourbonniere L, Larouche S, Terouz S, Bin JM, Lécuyer M-A, Saint-Laurent O, Larochelle C, Darlington PJ, Arbour N, Antel JP, Kennedy TE, Prat A (2015) Netrin 1 regulates blood–brain barrier function and neuroinflammation. Brain 138:1598–1612
32. Haqqani AS, Hill JJ, Mullen J, Stanimirovic DB (2011) Methods to study glycoproteins at the blood-brain barrier using mass spectrometry. Methods Mol Biol 686:337–353
33. Haqqani AS, Delaney CE, Brunette E, Baumann E, Farrington GK, Sisk W, Eldredge J, Ding W, Tremblay T-L, Stanimirovic DB (2018) Endosomal trafficking regulates receptor-mediated transcytosis of antibodies across the blood brain barrier. J Cereb Blood Flow Metab 38:727–740



Chapter 21

Genomics-Driven Immunoproteomics: An Integrative Platform to Uncover Important Biomarkers for Human Diseases

Raghavendra Giri, Veneta Qendro, Pooja Rani, Carren Jepchumba, Grace Bugos, Volker Stadler, and David K. Han

Abstract

Genomics-driven immunoproteomics (GDI) is a platform that helps identify antigenic protein targets of mutations and other deoxyribonucleic acid (DNA) variations that are commonly associated with pathological states. This platform utilizes data generated from deep sequencing of exomic DNA or ribonucleic acid (RNA) as input to synthesize mutant peptides into microarrays, which then can be used to detect antigenic proteins that invoke immune response in patients. The technology has been used to detect antigenic targets of multiple sclerosis, an autoimmune disease [1], and cancer to identify mutant proteins that invoke immune response in breast cancer patients [2]. This technology has many potential applications to select genomic changes that are specifically recognized by the immune system in a rapid and efficient manner.

Key words Immunoproteomics, Biomarker, Cancer diagnostic test, Autoimmune antigens, Cancer antigens, Neo-antigens, Integrative platform

1 Introduction

Genomic sequencing has become highly accessible and affordable, with the advent of technologies commonly referred to as next-generation sequencing (NGS). As a result of this, a vast amount of data is generated, giving researchers insight into single nucleotide variants (SNVs), mutations, insertions, and deletions in many genomes associated with human diseases. These changes can potentially be translated to changes in protein structure and function, which may impact physiological balance at the cellular and systems levels, such as the immune system recognizing tumor antigens expressed and presented exclusively by mutated genomes of cancer cells.

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_21, © Springer Science+Business Media, LLC, part of Springer Nature 2019

Changes in protein sequence can lead to binding and presentation of mutated peptides (neoantigens) to Major Histocompatibility Complex (MHC) receptors, which can invoke an immunogenic response. MHC class I and II presentation of these mutant peptides should allow immune system to detect these neo-antigens, provided T- and B-cell responses are functioning normally. Lack of efficient detection of mutant peptides is one of the primary causes of cancer escape from immune surveillance. Although genomic sequencing can provide mutation profiles rapidly, identification of mutant peptides that can invoke immune response in patients cannot be easily accomplished without the Genomics driven immunoproteomics (GDI) technological platform.

Selecting vast amount of genomic data from cancer cells rapidly to identify tumor antigen is where the bottleneck lies. The GDI technology was developed to overcome this limitation. Specifically, the GDI technology allows selection of patient-specific neo-antigens that are recognized by the immune system, among the vast amount of genomic mutations on protein coding regions (exons). Although cancer and autoimmune diseases have been used as case studies, this technology has potential to positively contribute toward many human and animal diseases, ranging from autoimmunity to infectious diseases. GDI can be used to potentially identify antigenic targets of rapidly multiplying infectious microorganisms like Middle East respiratory syndrome (MERS), Ebola, chikungunya, and many more that had major outbreaks. It is noteworthy that no effective vaccines have been discovered yet for many of these infectious diseases, mainly due to the fact that antigenic epitopes of infectious microorganisms have yet to be identified. By sequencing the infectious microorganism genomes and converting the resulting protein sequence to peptide microarrays, GDI can be used to uncover the reactive antigens that are recognized by antibodies present in the test sample or patients' sera, followed by developing and testing vaccines utilizing antigenic epitopes. Also, GDI technology can be used to identify neutralizing antibodies targeting antigenic epitopes on infectious microorganisms present in the test subject.

GDI technology may also be useful for identifying pathogenic mechanisms of neurological disorders such as Alzheimer disease. In one recent study, Readhead et al. found evidence that links the activity of specific herpes viral species, such as HHV6A and HHV7, with molecular, genetic, clinical, and neuropathological aspects of Alzheimer disease. GDI can potentially help detect antibody-mediated response to viral proteomes, which in turn may be helpful to monitor Alzheimer disease development and progression [3].

This technology may also have a positive role in discovering antibody-mediated immune processes that influence pathogenesis of orphan diseases. For example, the exact cause of acute hemorrhagic

leukoencephalitis, a disease in which brief and intense inflammatory attacks the brain and spinal cord are attributed to debilitating conditions frequently affecting children, is not known. Another example is amyotrophic lateral sclerosis (ALS), a disease presenting with weakness and stiffness symptoms due to upper and lower motor neuron cell death. Although autoimmune process has been implicated, specific antigenic targets have not been yet identified. GDI technology can help to identify the antigens and reacting antibodies the same way it was demonstrated in a case study of multiple sclerosis [2].

GDI technology allows effective integration of large-scale genomics, proteomics, and immune readout and, therefore, may be useful to examine potential antigenic targets as well as isolate reacting antibodies relevant to diseases in which pathogenic mechanisms are not easily identifiable. However, in order to comprehensively understand critical disease-causing mechanisms, additional investigations may be necessary in concert with the GDI technology. Future technological advancements in molecular, imaging, and biomarker analysis followed by more comprehensive integration with GDI-like technologies will likely help identify precise disease mechanisms and help develop strategies toward prevention, diagnosis, and management.

2 Materials

Prepare all solutions using deionized distilled ultrapure water (prepared by purifying deionized water, to attain a sensitivity of 18 MΩ-cm at 25 °C) and analytical grade reagents. Store all reagents at 4 °C (unless indicated otherwise). Diligently follow all waste disposal regulations when disposing of waste. Peptide microarrays and incubation trays were ordered from PEPperPrint Company (PEPperPRINT GmbH, Heidelberg, Germany).

1. Standard buffer: 250 μL of Tween 20 (0.05%) to 500 mL of ice-cold Phosphate Buffered Saline (PBS), pH 7.4.
2. Blocking buffer: 5 g of bovine serum albumin (BSA) powder (Fisher BSA Fraction V) in PBS, pH 7.4.
3. Staining buffer: Add 5 mL of blocking buffer to 95 mL of standard buffer. Prepare staining buffer fresh before every experiment to avoid contamination.
4. Prepare dipping deionized water in a Coplin staining jar.
5. Primary antibodies: Serum or plasma from patients or from model animals should be stored at 4 °C, tested for appropriate dilution, and prevented from repeated freeze-thaw cycles (*see* **Note 1**).

6. Both the primary and secondary antibody (sera/plasma) should be diluted appropriately in staining buffer.

3 Methods

3.1 Creating Peptide Microarrays

1. Exomic data from normal and diseased regions from the same patient or animal model should be compared and variants determined using any common genomic data analysis platforms (*see* **Note 2**).
2. Select a 15-mer mutant and corresponding wild-type peptides (SNVs, insertions, deletions, and frameshift mutations), and send out to PEPperPRINT or other similar companies for synthesis of peptide microarrays. Store arrays as per manufacturer's instructions.

3.2 Staining Peptide Microarrays and Data Analysis

1. Handle microarray just by touching the edge of the chip, avoiding touching the surface that contains peptide spots (Fig. 1).
2. Place the microarray in incubation tray, such as PEPperPRINT, ppc.006, or comparable tray, carefully and tighten the screws evenly and slowly to avoid breaking the glass microarray slides (Fig. 2).
3. Place dummy microarrays if there are empty cavities in the staining (*see* **Note 3**).
4. Incubation can be done at room temperature or 4 °C shaking in orbital shaker at 140 rpm.
5. Incubate microarray in standard buffer for 15 min.
6. Aspirate the standard buffer, then add blocking buffer, and incubate for 30 min. (*see* **Notes 4** and **5**).

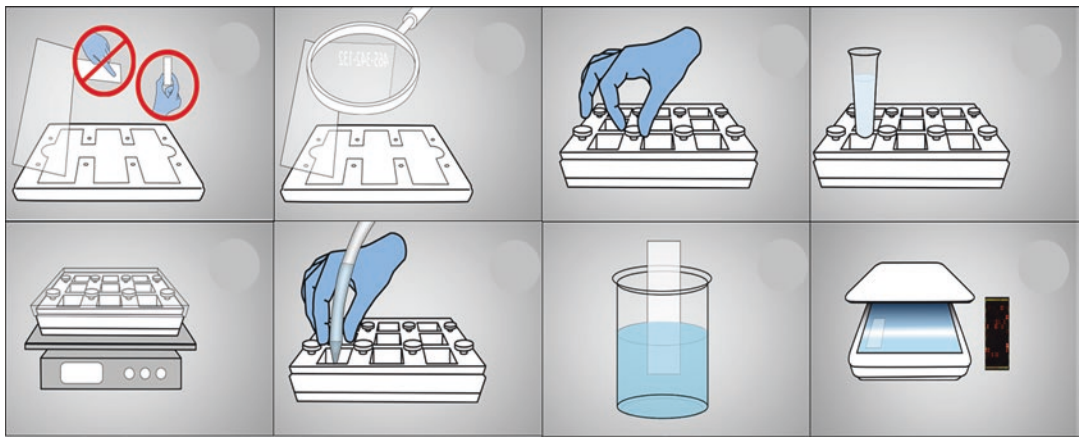


Fig. 1 Schematic drawing showing how to handle and incubate microarray tray and microarray slide

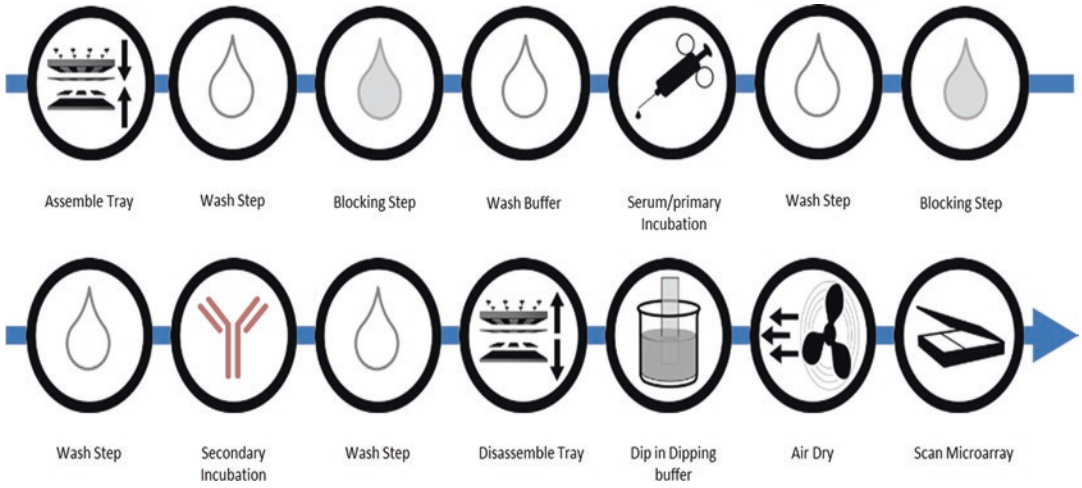


Fig. 2 Flow chart illustrating method to incubate microarray

7. Remove the blocking buffer, wash once with standard buffer, and then add primary sample.
8. Incubate shaking overnight at 4 °C.
9. Wash the microarray with standard buffer three times.
10. Incubate with secondary antibody at appropriate dilution (according to manufacturer) for 45 min.
11. Wash the microarray with standard buffer three times.
12. Disassemble microarray from incubation tray.
13. Dip in dipping buffer 2–3 times.
14. Air-dry then scan microarray.
15. Scan the peptide microarrays using GenePix 4000B scanner or equivalent microarray slide scanner.

3.3 Analysis of Data from GenePix

1. After the scanned image is saved in the folder of choice, load gal file.
2. Align grid with the image, just retaining the large and central grid.
3. From tool section on left side of screen, use block mode icon to move the grid to the image, and use zoom in and out to check the alignment.
4. Analyze the aligned image by checking the analyze icon on the right, just above the folder/save icon.
5. Spot-check the results, mouseover the image, and check their intensities in feature viewer (lower left side of screen).
6. Save result as gpr file and which could be further used for creation of excel file (peptide map, ratio table, etc.).

4 Notes

1. We had tested 1:10, 1:25, 1:50, and 1:500 dilutions of breast cancer patients' sera using peptide microarrays. We found that although the lowest dilution (1:10) and the highest dilution (1:500) showed similar qualitative results, the best quantification for mutant wild-type ratios can be achieved at 1:25 serum dilution. This might be different as per the serum being tested, so dilution should be tested as per need.
2. When comparing normal versus disease or binding of antibodies to wild-type versus mutant peptides, caution should be taken in assessing differential mutant binding affinity. We have validated that duplicate spots of each peptide are sufficient to generate high-confidence data [1]. Typically, duplicate set of wt-mut pairs were compared using the power law global error model (PLGEM) in order to identify statistically significant intensity differences. PLGEM software was downloaded from www.bioconductor.org and run in-house. A cutoff *p*-value of 0.01 was used as an initial filter for significance.
3. Placing dummy microarray will make sure the rubber gasket is seated properly and pressure is well distributed which prevents slides from breaking.
4. Solution should be added gently by placing the pipet tip at the corner of the array chamber without touching microarray surface.
5. Care should be taken while aspirating array chamber so as to avoid drying of the microarray surface during the process.
6. If signal intensities and/or signal-to-noise ratios are too low, the same array can be incubated again using your sample with a lower serum dilution or higher antibody concentration. Repeat incubation with your primary sample and staining with secondary antibody.
7. Keep the tray cover on while incubation with fluorescently label antibody to avoid photo bleaching.

References

1. Qendro V, Lundgren DH, Palczewski S et al (2017) Discovery of putative breast cancer antigens using an integrative platform of genomics-driven immunoproteomics. Proteomics 17:1600318
2. Qendro V, Bugos GA, Lundgren DH et al (2017) Integrative proteomics, genomics, and translational immunology approaches reveal mutated forms of Proteolipid Protein 1 (PLP1) and mutant-specific immune response in multiple sclerosis. Proteomics 17:1600322
3. Readhead B, Haure-Mirande J-V, Funk CC et al (2018) Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. Neuron 99:64–82.e7



Chapter 22

Identification of Antigens in Immune Complexes

Nozomi Aibara and Kaname Ohyama

Abstract

Comprehensive identification and profiling of antigens in immune complexes (IC-antigens) is useful to provide insights into pathophysiology and could form the basis for novel diagnostic and treatment strategies for many immune-related diseases. Immune complexome analysis is the method for comprehensively identifying and profiling IC-antigens in biological fluids (such as serum and cerebrospinal fluid). Here, we describe an IC-antigen detection method; specifically, ICs in biological fluids are captured by using protein G- or protein A-coated beads, are subjected to papain-digestion, elution, and tryptic digestion, and are analyzed by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS).

Key words Antibody, Antigen, Immune complex, Nano-LC-MS/MS, Papain

1 Introduction

Immune complexes (ICs) are formed upon noncovalent interaction between foreign antigens or autoantigens and antibody proteins [1]. Antigens in ICs (IC-antigens) are targets of the immune system, and an increase of some ICs leads to autoimmune diseases, infections, and cancers [2]. Therefore, comprehensive identification and profiling of IC-antigens is useful to provide insights into pathophysiology and could form the basis for novel diagnostic and treatment strategies for these diseases. However, such studies have been limited to date because tools for screening of ICs are lacking.

We have developed a method, designated "immune complexome analysis," to catalogue antigens in ICs [3, 4] and have used this method to identify specific antigens in circulating ICs in serum for autoimmune diseases [3, 5–7], infectious diseases [8], and cancers [9]. Many proteomics methods for immunological studies focus on free antigens or antibodies. In contrast, our method focuses on IC-antigens among all the proteins contained in the human body. In our method, ICs are isolated from biological fluids, such as serum and cerebrospinal fluid, by using protein G- or protein A-coated beads that bind the fragment of crystallization

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_22, © Springer Science+Business Media, LLC, part of Springer Nature 2019

domain of antibodies. The ICs are then subjected to papain-digestion and elution. Papain selectively cleaves immunoglobulin (Ig) at the heavy chain hinge region into three fragments, one Fc and two identical Fab fragments [10–12], resulting in selective dissociation of the antigens from ICs without eluting the proteins bound non-specifically to the beads. The eluted IC-antigens are tryptically digested, and the resulting peptide digests are analyzed using nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS). Figure 1 summarizes immune complexome analysis for identification of IC-antigens.

2 Materials

2.1 Pretreatment of Clinical Samples

1. Biological fluid (e.g., serum, urine, cerebrospinal fluid) (*see* **Note 1**).
2. Protein A- or protein G-coated magnetic beads (*see* **Note 2**).
3. Phosphate-buffered saline (PBS): 9.0 mM Na_2HPO_4, 2.9 mM NaH_2PO_4, and 137 mM NaCl, pH 7.4.
4. Papain: 0.01 mg/mL in 40 mM ethylenediaminetetraacetic acid (EDTA) and 40 mM L-cysteine (*see* **Note 3**).
5. Dithiothreitol (DTT):10 mM in water.
6. Iodoacetamide (IAA): 55 mM in water.
7. Ultrapure water.

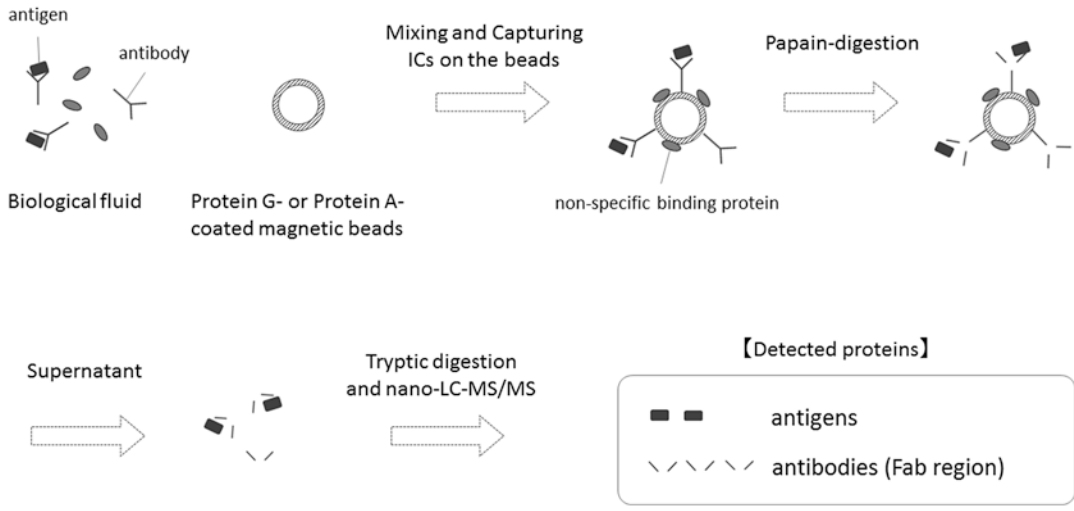


Fig. 1 Summary of the IC-antigen detection process by immune complexome analysis. ICs are captured from biological fluids by using protein G- or protein A-coated beads. The ICs are then subjected to papain-digestion and elution. The eluted IC-antigens are tryptically digested and are analyzed using nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS)

8. Acetic acid: 0.05% (v/v) in water.
9. Ammonium hydrogen carbonate: 50 mM in water.
10. Trypsin: 0.5 g/L.
11. Trifluoroacetic acid (TFA): 10% (v/v) in water.
12. Water bath.
13. Vortex mixer.
14. Stirring centrifuge.
15. Centrifugal evaporator.

2.2 Protein Identification by Nano-LC-MS/MS

1. Mobile phase A: ultrapure water and 0.2% (v/v) formic acid.
2. Mobile phase B: 0.1% (v/v) formic acid in 90% (v/v) acetonitrile.
3. Nano-precolumn for desalting enrichment.
4. Nano-liquid chromatography (LC) column for separating the digests (C18, 75 μm i.d. × 125 mm, 3 μM particle, 100 Å pore size, Nikkyo Technos).
5. Nano-LC system.
6. Mass spectrometry (MS) system.
7. Protein identification software (e.g., Proteome Discoverer (Thermo Fisher Scientific)).
8. Protein database: UniProt Knowledgebase (human).
9. Bovine serum albumin (*see* **Note 4**).
10. HeLa Protein Digest Standard (Thermo Fisher Scientific) (*see* **Note 4**).

3 Methods

3.1 Pretreatment of Clinical Samples

1. Pipette the protein G- or protein A-coated magnetic beads (40 μL) into a 1.5 mL of microtube, and wash the beads by adding 500 μL of PBS and vortexing for 10 s.
2. Place the tube into the magnetic stand, allow the beads to migrate to the magnet, and then remove the washing buffer with a pipette.
3. Incubate the beads with 10 μL of human serum diluted with 90 μL of PBS for 30 min with gentle mixing, and capture ICs on the beads (*see* **Note 5**).
4. Place the tube into the magnetic stand, and then remove the supernatant with a pipette.
5. Wash the beads three times with 500 μL PBS (place the tube into the magnetic stand, allow the beads to migrate to the magnet, and then remove the washing buffer with a pipette).

6. Suspend washed beads in 50 μL of papain solution (40 mM EDTA, 40 mM L-cysteine), and incubate at 37 °C for 30 min (*see* **Note 3**).
7. Transfer the supernatant to another tube with a pipette (*see* **Note 6**).
8. Add 50 μL of 55 mM IAA dissolved in PBS to the sample (*see* **Note 7**).
9. Add 100 μL of 10 mM DTT to the sample, and further incubate it at 56 °C for 45 min (*see* **Note 8**).
10. Add 100 μL of 55 mM IAA to the sample, and incubate it at room temperature in the dark for 30 min (*see* **Note 8**).
11. Add 100 μL of 50 mM ammonium hydrogen carbonate and 100 μL of ultrapure to the sample (*see* **Note 9**).
12. Add trypsin in 0.05% acetic acid (final concentration of 0.5 g of trypsin/L) to the sample, and incubate it overnight at 37 °C (*see* **Note 9**).
13. Add 12 μL of 10% TFA to the sample, and mix it for 2 min to stop the digestion.
14. Vacuum-reduce the sample (about 500 μL) to approximately 80 μL, and store at 4 °C for subsequent analysis by nano-LC-MS/MS.

3.2 Protein Identification by Nano-LC-MS/MS

Nano-LC-MS/MS analysis of final digest sample (1 μL) (*see* **Note 10**) obtained by the pretreatment procedure as follows:

1. The separation is performed by using the mobile phase A (0.1% formic acid) and mobile phase B (0.1% formic acid in 90% acetonitrile) with a gradient elution (e.g., 5% to 33% mobile phase B in 20 min and 100% mobile phase B held for 5 min).
2. Ion-spray into MS (e.g., Q Exactive) with a spray voltage (e.g., 1.5 kV).
3. Collect spectra with a mass/charge ratio range of 400–1500.
4. The transfer capillary temperature is set at 200 °C.
5. Extract the MS/MS data using protein identification software (e.g., Proteome Discoverer).
6. Search the spectra against sub-databases from the public non-redundant protein database of UniProt Knowledgebase (human) with the following search parameters: mass type, monoisotopic precursor, and fragments; enzyme, trypsin (KR); enzyme limits, full enzymatic cleavage allowing up to two missed cleavages; ion and ion series calculated, B and Y ions; static modification, C (carbamidomethylation); and dif-

ferential modifications, M (oxidation), N, and Q (deamidation) (search engine, e.g., SEQUEST).

7. The filter criteria (single, double, and triple charge peptides with a correlation factor [XCorr] and protein probability [P]) are adjusted maintaining the empirically determined protein false discovery rate at 5%.
8. Data processing is performed in Excel to identify specific IC-antigens (*see* **Note 11**).

4 Notes

1. This method is suitable for biological fluids which contains immune complexes, for example, serum, urine, and cerebrospinal fluid.
2. We recommend using magnetic beads because it is suitable for washing steps and isolating supernatant. Protein G is known to capture all the types of IgG, and protein A is known to capture IgA, IgD, IgE, IgM, IgG1, IgG2, and IgG4. These differences of affinity enable to comprehensively detect CIC-antigens and a little overlap between protein G and protein A in immune complexome analysis.
3. The concentration of papain solution is based on the method used.
4. Check the performance of the nano-LC-MS/MS system by confirming the peak shapes of fully digested peptides (20 fmol) derived from bovine serum albumin and the number of the proteins from HeLa Protein Digest Standard (100 ng/μL) at the beginning of the measurement.
5. Adjust the amount of biological fluids mixed with magnetic beads in response to the concentration of protein contained in each biological fluid.
6. Transfer the supernatant to a new tube without aspirating beads.
7. Stop the papain-digestion reaction by adding iodoacetamide to the sample. This is based on the method selected [13].
8. Proteins are reduced by DTT and alkylated by IAA in order to facilitate trypsin digestion.
9. Since the optimal pH of trypsin is around 8–9, the pH is adjusted with ammonium hydrogen carbonate.
10. The volume of injection into the LC system can be changed in response to protein concentration.
11. The proteins except for immunoglobulin, complements, and bead-derived proteins are listed for each disease with their detection frequencies (Fig. 2).

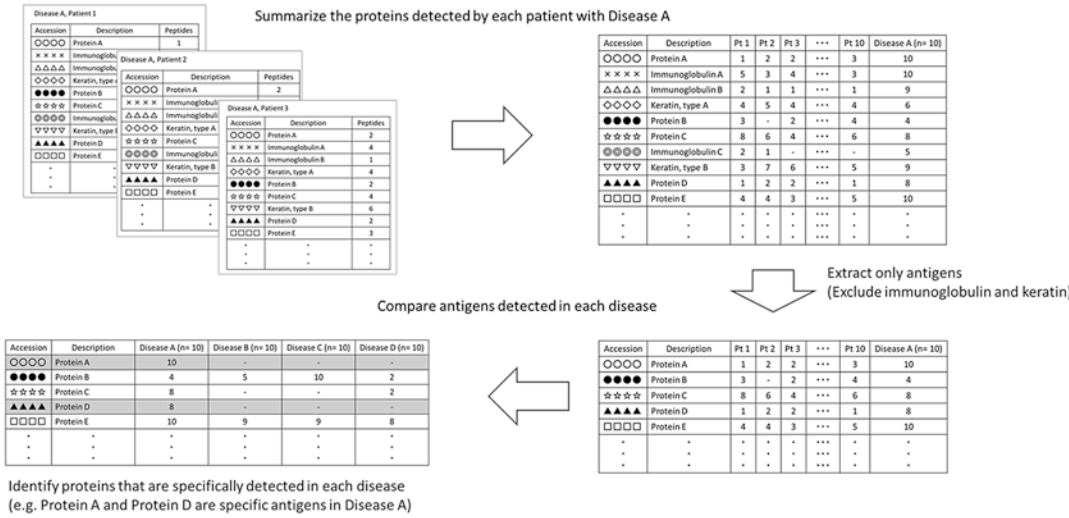


Fig. 2 Illustration of the process for listing IC-antigens in each disease and for finding disease-specific IC-antigens

References

1. Nezlin R (2000) A quantitative approach to the determination of antigen in immune complexes. J Immunol Methods 237:1–17
2. Chauhan AK (2017) Editorial: immune complexes in disease pathology. Front Immunol 8:173
3. Ohyama K, Ueki Y, Kawakami A, Kishikawa N, Tamai M, Osaki M, Kamihira S, Nakashima K, Kuroda N (2011) Immune complexome analysis of serum and its application in screening for immune complex antigens in rheumatoid arthritis. Clin Chem 57:905–909
4. Baba M, Ohyama K, Kishikawa N, Kuroda N (2013) Optimization of separation and digestion conditions in immune complexome analysis. Anal Biochem 443:181–186
5. Ohyama K, Kawakami A, Tamai M, Baba M, Kishikawa N, Kuroda N (2012) Serum immune complex containing thrombospondin-1: a novel biomarker for early rheumatoid arthritis. Ann Rheum Dis 71:1916–1917
6. Ohyama K, Baba M, Tamai M, Yamamoto M, Ichinose K, Kishikawa N, Takahashi H, Kawakami A, Kuroda N (2015) Immune complexome analysis of antigens in circulating immune complexes isolated from patients with IgG4-related dacryoadenitis and/or sialadenitis. Mod Rheumatol 26:248–250
7. Ohyama K, Baba M, Tamai M, Aibara N, Ichinose K, Kishikawa N, Kawakami A, Kuroda N (2015) Proteomic profiling of antigens in circulating immune complexes associated with each of seven autoimmune diseases. Clin Biochem 48:181–185
8. Ohyama K, Huy NT, Yoshimi H, Kishikawa N, Nishizawa JE, Roca Y, Revollo Guzmán RJ, Velarde FU, Kuroda N, Hirayama K (2016) Proteomic profile of circulating immune complexes in chronic Chagas disease. Parasite Immunol 38:609–617
9. Ohyama K, Yoshimi H, Aibara N, Nakamura Y, Miyata Y, Sakai H, Fujita F, Imaizumi Y, Chauhan AK, Kishikawa N, Kuroda N (2017) Immune complexome analysis reveals the specific and frequent presence of immune complex antigens in lung cancer patients: a pilot study. Int J Cancer 140:370–380
10. Bennett KL, Smith SV, Truscott RJ, Sheil MM (1997) Monitoring papain digestion of a monoclonal antibody by electrospray ionization mass spectrometry. Anal Biochem 245:17–27
11. Moorhouse KG, Nashabeh W, Deveney J, Bjork NS, Mulkerrin MG, Ryskamp T (1997) Validation of an HPLC method for the analysis of the charge heterogeneity of the recombinant monoclonal antibody IDEC-C2B8 after papain digestion. J Pharm Biomed Anal 16:593–603
12. Adamczyk M, Gebler JC, Wu J (2000) Papain digestion of different mouse IgG subclasses as studied by electrospray mass spectrometry. J Immunol Methods 237:95–104
13. Andrew SM, Titus JA (2001) Fragmentation of immunoglobulin G. Curr Protoc Immunol Chapter 2:Unit 2.8



Chapter 23

Analysis and Characterization of Immune Cells and Their Activation Status by Whole-Cell MALDI-TOF Mass Spectrometry

Richard Ouedraogo, Julien Textoris, Laurent Gorvel, Aurélie Daumas, Christian Capo, and Jean-Louis Mege

Abstract

For 40 years, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been widely used in proteomics and biochemistry. It has been demonstrated in the last decade that MALDI-TOF MS can be used routinely to identify and classify numerous bacterial species or subspecies. We applied MALDI-TOF MS directly to intact mammalian cells, and we found that this method is valuable to identify human circulating cells and cells involved in the immune response including macrophages. We then stimulated human macrophages with cytokines, bacterial products, and a variety of bacteria. We found that MALDI-TOF MS discriminated unstimulated and stimulated macrophages and also detected multifaceted activation of macrophages. We conclude that whole-cell MALDI-TOF MS is an accurate method to identify various cell types and to detect subtle modifications in cell activity and therefore it can be beneficial in clinical practices for a rapid patient classification based on their immune profile.

Key words Mass spectrometry, Matrix-assisted laser desorption/ionization time-of-flight, Intact cell, Macrophage, Cell activation

1 Introduction

Since the 1980s, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been widely utilized in biochemistry. It was not until the last decade that MALDI-TOF MS was exploited in microbiology and clinical microbiology labs. MALDI-TOF MS was demonstrated to be a rapid and easy to use tool for the identification of microorganisms [1]. The ability of the technique to identify whole-cell microorganisms to the level of species and subspecies has led to a gain in popularity of the approach. In practice, identification is based on the generation of mass spectra from the entire microorganism and comparing these spectra with a database of reference spectra, to

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_23, © Springer Science+Business Media, LLC, part of Springer Nature 2019

provide results in minutes. Thus the MALDI-TOF has been used routinely [2] for the identification of microbes.

There was interest in applying this approach for the profiling of mammalian cells. However, sample preparation steps, including ethanol or formic acid to extract the cell proteins, did not provide good quality MS spectra. Traditional use of MALDI-TOF MS methods used in proteomics research is not yet adapted to clinical small samples studies. The majority of clinical immunological sample are cells, plasma, or serum. It does require several millions of cells from biological sample to be sufficient for classical proteomics studies. In this case, it will be very difficult to utilize the proteomic tools to study cell activation at the individual level. It also requires a good practical knowledge of proteomics procedures. In addition, biomarker research using patient serum is demanding. It goes through several steps; from the concentration of the proteins, migration of the samples, and separation on polyacrylamide gels. Then prepare the proteins for digestion and for MALDI-TOF MS/MS analysis while avoiding contamination with keratin. Finally, it is necessary to have knowledge of computer science and bioinformatics to analyze the peptide sequences, to assemble them, and to predict their coverage in the sequence of a given protein [3, 4]. In other words, all of its requirements make it almost impossible to use classical proteomics as routine diagnostic tools in clinic.

Our method makes it possible to overcome the difficulties mentioned above and to increase the use of MALDI-TOF in clinical intact cells studies. Indeed, it makes it possible to measure the state of global activation of the immune cells of circulating blood or macrophages derived from monocyte activation [5]. In addition, macrophages have a high degree of plasticity and adapt quickly in response to their microenvironment. They sense microorganisms through receptors that bind conserved and ubiquitous microbial motifs, such as lipopolysaccharide (LPS). Macrophages stimulated with type I cytokines, i.e., IFN-γ, tumor necrosis factor (TNF), or bacterial products, i.e., LPS, adopt an M1 macrophage phenotype which is inflammatory, tumoricidal, and microbicidal. Macrophages stimulated with interleukin (IL)-4 or IL-10 adopt an M2 macrophage phenotype, which regulates inflammatory and immune responses and is only weakly tumoricidal and microbicidal [6]. Hence, macrophages constitute a model of choice to assess the accuracy of whole-cell MALDI-TOF MS to detect subtle modifications in cell activity [7].

We describe here the effective use of MALDI-TOF MS to identify several intact eukaryotic cell populations through the creation of a database of known samples. The experimental protocol, bioinformatics analysis of whole-cell MALDI-TOF MS spectra, and the comparison of unknown samples to the database allow the identification of various cell types within heterogeneous samples

[5] or multifaceted activation of macrophages [7]. Our method has the potential to be used for screening of patient-derived samples in order to profile cells according to the state of activation of their immune system.

2 Materials

The different cell types or stimulated cells were prepared separately to build databases. Experiments with mixed cell types (circulating cells) are indicated. Prepare sterile solutions for cell isolation and culture. Prepare and store all reagents at 4 °C. Prepare MALDI-TOF matrix just before use.

2.1 Human Circulating Cells

1. Isolate peripheral blood mononuclear cells (PBMCs) from leukopacks (or blood donors) by Ficoll gradient, as previously described [5].
2. Prepare $CD14^+$ monocytes using CD14 MicroBeads and the MACS separation system (Miltenyi Biotec), according to the manufacturer's protocol.
3. Prepare T $CD3^+$ lymphocytes using CD3 MicroBeads and the MACS separation system, according to the manufacturer's protocol.
4. Monocyte-derived macrophages (MDMs), obtained by incubating monocytes (10^6 cells in 6-well plates) in 3 mL of RPMI 1640 containing 20 mM HEPES, 10% human serum AB^+, 2 mM L-glutamine, 100 IU/mL penicillin, and 100 μg/mL streptomycin for 4 days. Replace human serum with fetal calf serum (FCS) for 3 additional days. The obtained cell population was identified as macrophages (more than 95%) by acquisition of membrane CD68 expression and CD14 down-modulation.
5. Dendritic cells (DCs), derived by incubating monocytes with 1000 U/mL human recombinant *granulocyte macrophage-colony stimulating factor* (GM-CSF) and 500 U/mL of human recombinant IL-4 in RPMI 1640 containing 10% FCS, 2 mM L-glutamine, 100 IU/mL penicillin, and 50 μg/mL streptomycin for 7 days. Change the medium every 3 days, and add again 1000 U/mL GM-CSF and 500 U/mL IL-4. The obtained cells expressed high levels of CD11c and CD1a and low levels of CD14 and CD68.
6. Polymorphonuclear cells (PMNs) obtained after Ficoll centrifugation, by sedimentation of red blood cells (RBCs) within dextran T500 (1.5% (w/v)) (*see* **Note 1**).
7. Obtain red blood cells (RBCs) by 1/1000 dilution of blood in phosphate buffered saline (PBS).

2.2 Noncirculating Cells

The preparation of several cell populations including mammalian primary cells and cell lines, one *Xenopus laevis* cell line (XTC-2 cells), and four types of amoebae (*Acanthamoeba polyphaga*, *Acanthamoeba castellanii*, *Hartmannella vermiformis*, *Poteriochromonas melhamensis*) is described in reference [5].

2.3 Macrophage Stimulation

1. Stimulate MDMs with 20 ng/mL of human recombinant IFN-γ, IL-4, IL-10, or TNF (purchased from R&D Systems or other suppliers) for different time points. Also stimulate MDMs with 1 μg/mL LPS from *Escherichia coli* or heat-killed bacteria (50 bacteria per cell) including *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium avium*, *Rickettsia prowazekii*, and *Orientia tsutsugamushi* (*see* **Note 2**).
2. *M. tuberculosis* (CIP H37Rv strain 103471), *M. bovis* (Bacillus Calmette-Guérin, BCG CIP strain 671203), and a clinical isolate of *M. avium* subsp. *hominissuis* were obtained from the Laboratory of Microbiology of the Hospital La Timone (Marseille) [6].
3. *R. prowazekii* strain Breinl (ATCC VR-142), the agent of epidemic typhus [8], and *O. tsutsugamushi* strain Kato (CSUR R163) [9], the agent of scrub typhus, were propagated in L929 cells (*see* **Note 3**).
4. Heat-killed bacteria were obtained after heating microorganisms at 95 °C for 1 h.

2.4 Reagents and Materials for MALDI-TOF MS

1. The matrix solution consists of a 10 mg/mL solution of acid α-cyano-4-hydroxycinnamic (HCCA) diluted in 50% acetonitrile, 25% Milli-Q grade water, and 25% trifluoroacetic acid at 10%. Mix and sonicate for at least 20 min. Centrifuge at 13,000 × *g* for 5 min. Discard the pellet and keep the supernatant. The matrix solution is ready for use (*see* **Note 4**).
2. AutoFlex II mass spectrometer (Bruker Daltonics).
3. Bruker MSP 384 software polished steel target (Bruker Daltonics).
4. FlexControl 3.0 software (Bruker Daltonics).
5. FlexAnalysis 3.3 software (Bruker Daltonics).
6. Biotyper 3.0 software (Bruker Daltonics).
7. ClinProTools 2.2 software (Bruker Daltonics).

3 Methods

3.1 MALDI-TOF Target Preparation

1. Moisten the Bruker MSP 384 polished steel target with hot tap water. Rub with KIMTECH paper. Rub with 70% ethanol. Rinse with hot tap water by rubbing with KIMTECH paper. Rub finally with 70% ethanol.

2. Immerse the target in 70% ethanol and sonicate for at least 15 min.
3. Cover the target with 500 µL to 1 mL of trifluoroacetic acid. Rub with KIMTECH paper. Rinse with HPLC water without rubbing. Dry the target at room temperature (*see* **Note 5**).

3.2 Preparation of Deposits

1. Centrifuge cells (2×10^6 cells per assay) at $300 \times g$ for 5 min, and wash them in PBS without Ca^{2+} or Mg^{2+} (*see* **Note 6**). Centrifuge to remove traces of culture medium. Collect cell pellets in 20 µL of PBS without Ca^{2+} or Mg^{2+}. Freeze cells for 2–3 days before analysis at −80 °C.
2. Thaw samples gently on ice (4 °C) (*see* **Note 7**).
3. Put the Bruker MSP 384 polished steel target on a horizontal support to obtain uniform deposits throughout the spot (*see* **Note 8**).
4. Homogenize cells in Eppendorf tubes before deposition of 1 µL on the MALDI target. Add 1 µL of the HCCA matrix to the sample on the target. Avoid mixing the spot with the pipette (*see* **Note 9**).
5. Drop 12–16 different spots of the same sample.
6. The evaporation takes place gradually at room temperature, and the formed HCCA crystals contain dispersed sample molecules.
7. Samples may be immediately analyzed or store in the dark for several days before analysis.

3.3 Acquisition of Data

1. Insert the Bruker MSP 384 polished steel target containing samples in the Autoflex II mass spectrometer (Bruker Daltonics) outfitted with the Compass 1.2 software suite (consisting of FlexControl 3.0 and FlexAnalysis 3.3 from Bruker Daltonics). Run samples in positive mode, with 240 laser satisfactory shots in 40 shot steps intervals and 40% laser power, performed in different regions of the analyzed sample spot. A signal-to-noise ratio of 3.0 was selected to define peaks, with a maximum of 100 peaks per spectrum. After the target plate calibration was complete, the AutoExecute command was used to analyze the samples. The processes described below are manufacturer or software defaults, and do not require adjusting.
2. Laser settings. Fuzzy control, on; weight, 2.00; laser power, between 30% and 45%; matrix blaster, 5.
3. Data evaluation. Peak selection masses, 4000–10,000; mass control list, off; peak exclusion, ignore the largest peaks in the defined mass range; peak evaluation processing method, default; smoothing, on; baseline subtraction, on; peak, resolution higher

than 400; protein maximal resolution, ten times above the threshold.

4. Accumulation. Parent mode, on; sum up to 240 satisfactory shots in 40 shot steps; dynamic termination, on; criteria intensity, early termination if reaching intensity value of 20,000 for 10 peaks.
5. Movement. Spiral large, maximal allowed number at one raster position; parent mode, 80; quit sample after 25 subsequently failed judgments.
6. Processing. Flex analysis. Method, none; BioTools MS method, none.
7. Sample carrier. Manual fine control speed, *x* (10,000 μm/s) *y* (20,000 μm/s); relative *x* (5 μm) *y* (5 μm); absolute *x* (−2000 μm) *y* (−2000 μm); state *x* (0) *y* (0); random walk, 50 shots at raster spot; mode, off.
8. Spectrometer. High voltage, switched on; ion source 1, 19.99 kV; ion source 2, 18.74 kV; lens, 7 kV; pulsed ion extraction, 330 ns; polarity, positive; matrix suppression, mode gaiting; gaiting strength, height; suppress up to, *m/z* 1500 Da.
9. Detection. Mass range, 2000–20,137; mode, medium range; detector gain, linear (18*x*); sample rate, 1.00 GS/s; electronic gain, enhanced (100 mV); real-time smooth, high; spectrum, size (63,463 pts), delay (29,412 pts).
10. Processing method. MBT process.
11. Setup. Range, medium; laser frequency, 25 Hz; autoteaching, on; instrument-specific settings, digitizer trigger level (2000 mV), digital off linear (127 cnt), digital off reflector (127 cnt); detector gain voltage offset, linear (1300 V), reflector, 1400 V; laser attenuator, offset (75%), range (15%); electronic gain button definitions, gain, regular (Offset Lin, 100 mV; Offset Ref 100 mV; 200 mV/full scale); enh. (Offset Lin, 51 mV; Offset Ref 51 mV; 100 mV/full scale); highest (Offset Lin, 25 mV; Offset Ref 25 mV; 50 mV/full scale).
12. Calibration. Calibration strategy, interactive; mass control list, bacterial test standard; zoom range, ±5%; peak assignment tolerance, user defined (1000 ppm); mode, linear.

3.4 Data Analysis

3.4.1 Spectrum Analysis

1. The FlexAnalysis software 3.3 allows raw spectrum processing, baseline subtraction, smoothing, and peak list editing and displays several spectra into one window or superimposes spectra (for the comparison of different types of circulating cells, *see* Fig. 1).
2. The ClinProTools 2.2 software from Bruker Daltonics is used to analyze the variability between different samples. Load spectra of each cell category to create according classes.

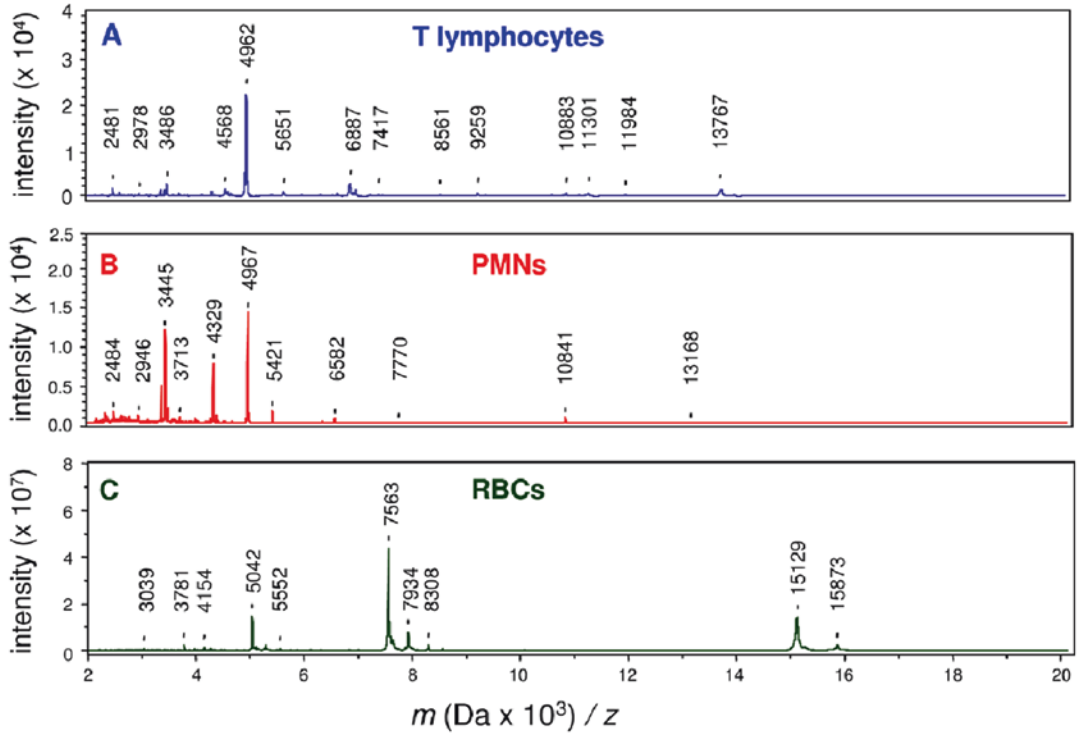


Fig. 1 MALDI-TOF MS spectra of circulating cells. T lymphocytes (**a**), PMNs (**b**), and RBCs (**c**) were isolated from a healthy blood donor. Representative MALDI-TOF MS spectra are shown. The figure is extracted from the reference [5]

2D representation generated by ClinProTools 2.2. The software selects automatically two peaks that are present in each cell type but have different intensities (here, the peaks 2 and 30). This representation highlights the reproducibility of spectra between spectra of each class (here, ten spectra obtained from ten different blood donors) and the differences between the different classes (here, monocytes, T lymphocytes, and PMNs) (Fig. 2).

3. Gel view representation generated by ClinProTools 2.2. This representation compares the reproducibility of spectra within the same class and the differences between different classes. The different bands represent different peaks of each class. The intensity of bands corresponds to the intensity of detected peaks. This representation shows the reproducibility of spectra within each class (here, four spectra obtained from four different blood donors) and the peaks that are differentially expressed in different classes (here, monocytes and T lymphocytes isolated from each blood donor) (Fig. 3).

3.4.2 Database Creation

1. The Biotyper 3.0 software from Bruker Daltonics is used to create and manage databases. An averaged spectrum for each cell category corresponds to at least 10 individual spectra. Here, we

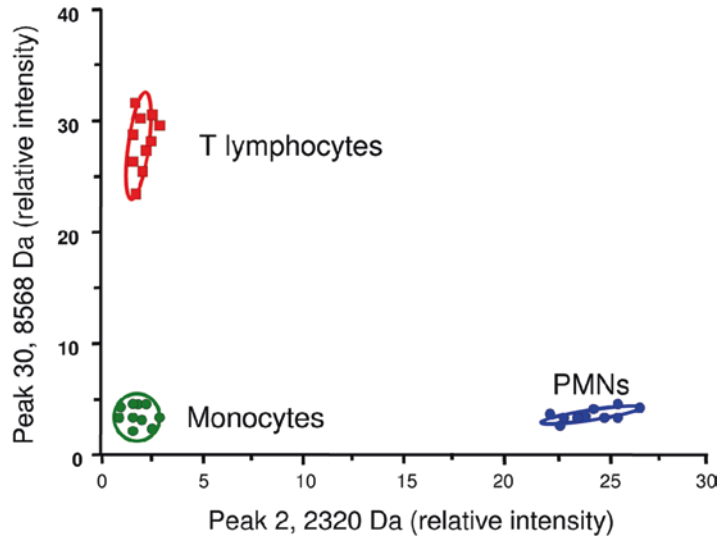


Fig. 2 Reproducibility of MALDI-TOF MS signatures. Monocytes, T lymphocytes, and PMNs were isolated from ten healthy blood donors. MALDI-TOF MS spectra were analyzed using 2D peak distribution view. The relative intensities of the two peaks automatically selected were homogenous among blood donors, and the ellipses represent the standard deviation within each cell population (monocytes, T lymphocytes, and PMNs, respectively). *See* reference [5]

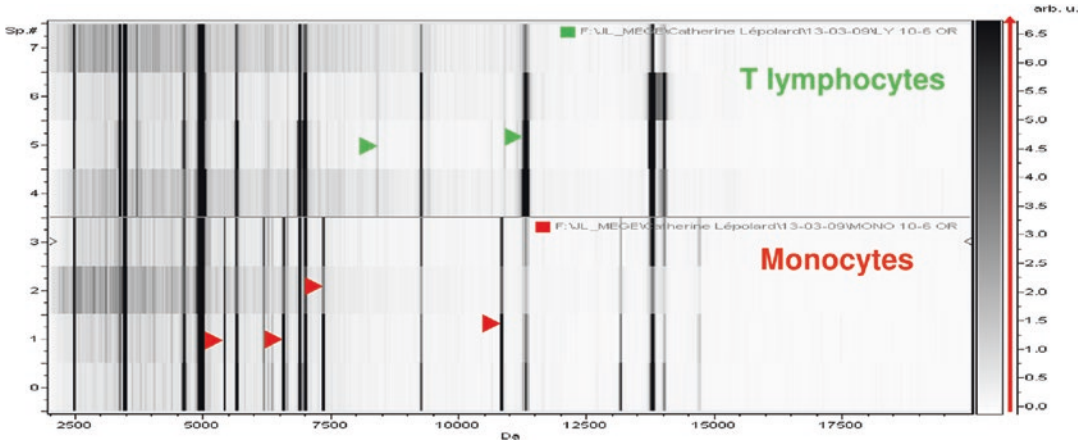


Fig. 3 Gel view representation of monocytes and T lymphocytes. Circulating cells were isolated from four different healthy blood donors. MALDI-TOF MS spectra are presented in gel view representation. Spectra are shown with *m*/*z* values on the x-axis, and the peak intensity (in arbitrary units) is coded with the gray scale presented on the right. Major differences between monocytes and T lymphocytes are indicated by arrowheads. *See* reference [5]

created a database that includes 17 mammalian cell types, 1 *X. laevis* cell line (XTC-2 cells), and 4 types of amoebae (*A. polyphaga*, *A. castellanii*, *H. vermiformis*, *P. melhamensis*) (*see* **Note 10**).

2. Baselines are automatically subtracted from spectra and the background noise smoothed. An average spectrum is automatically created using default Biotyper method settings provided by the manufacturer. The sensitivity (the maximum tolerated error) of mass spectrum values and spectrum shifts was 200 particles per million. The minimum frequency to benchmark selection of peaks was 25%, and only peaks with signal/noise intensity above background are automatically selected by the software. An average virtual spectrum consisting of 70 peaks is added to the database as a new reference.
3. The Biotyper 3.0 software is used to generate a dendrogram representation of cell categories according to their protein fingerprint (Fig. 4). This figure shows that mammalian cells (in red and green) and nonmammalian cells (in black and blue) were in two distinct branches of the dendrogram. Circulating nucleated cells (in green) clustered within a sub-branch distinct of primary cells (trophoblasts, DCs, MDMs, murine bone marrow-derived macrophages (BMDMs)) and cell lines. Note that human RBCs, which are nonnucleated cells, clustered with nonmammalian cells.
4. The Biotyper 3.0 software is used to identify unknown spectra by comparison with database references. The sensitivity and the maximum error tolerated to determine the mean m/z values are 1 Da. Score values between 0.000 and 1.699 indicate that the unknown spectra did not match with known references. Values between 1.700 and 1.899 indicate probable cell identification. Scores between 2.000 and 3.000 are considered statistically significant and allow effective identification of the unknown spectra [5]. This procedure is currently used to identify bacterial species in clinical samples [8]. Here, we extend this method and the scores provided by Bruker Daltonics to identify the different cell populations present in a complex tissue. First, isolated monocytes and T lymphocytes were mixed, and the resulting fingerprint shows that monocytes and T lymphocytes are identified by MALDI-TOF MS (Table 1). Table 1 also shows that the fingerprints of monocytes and T lymphocytes, respectively, are identified in peripheral blood mononuclear cells. In whole blood that contains leukocytes essentially composed of PMNs, the fingerprint of PMNs is identified, but not those of monocytes and T lymphocytes.

3.5 Macrophage Activation Analysis with R Software

The analysis of MALDI-TOF MS spectra described above is performed using Bruker Daltonics software. We present here a similar analysis performed with an open-source software (R).

1. Load raw spectra in R (version 2.14) using the readBrukerFlexData library.

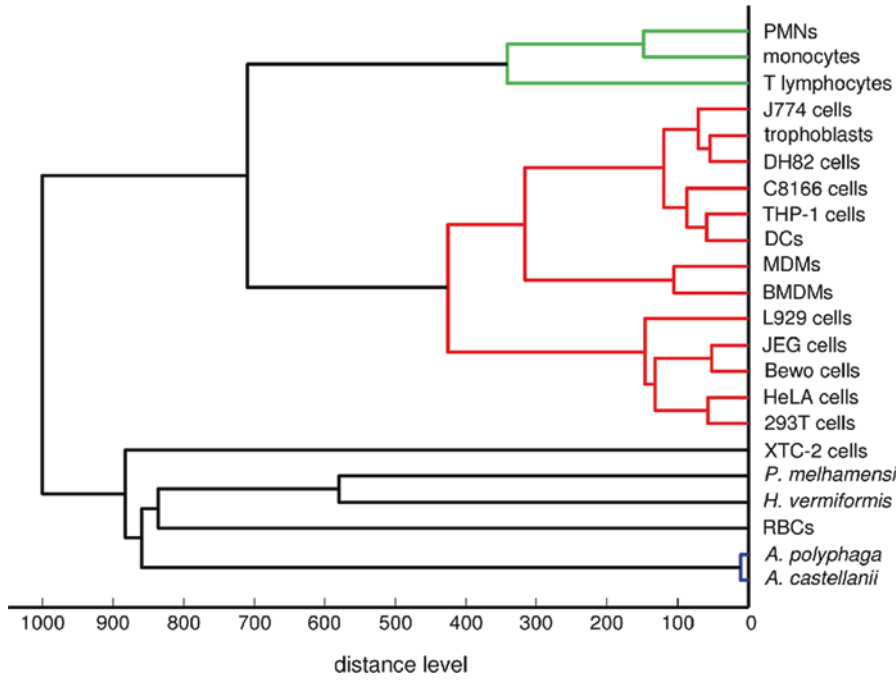


Fig. 4 Dendrogram of 22 eukaryotic cell types. MALDI-TOF MS was performed on 22 cell types with at least 20 spectra per cell type. A mean spectrum for each cell type was added to the database using the BioTyper 3.0 software and the dendrogram creation method. *See* reference [5]

Table 1
Identification of subpopulations by MALDI-TOF MS. The fingerprints of monocytes and T lymphocytes were identified in a mixed population (50% monocytes, 50% T lymphocytes) and in PBMCs. In blood, it was possible to identify PMNs (that represent about 70% of total leukocytes), but not monocytes or T lymphocytes

Multiple cell population	Identification	Scores
Monocytes + T lymphocytes (equal concentration)	Monocytes T lymphocytes	2250 2247
PBMCs	Monocytes T lymphocytes	2078 2024
Whole blood (after hypotonic shock)	PMNs Monocytes T lymphocytes	2049 1585 1654

2. Analyze spectra using the MALDIquant library and specific algorithms. The square root of the intensities is used to enhance graphical visualization of the spectra.
3. Correct background using statistics-sensitive nonlinear peak-clipping algorithm for baseline estimation [10]. Peaks are

detected using a signal-to-noise ratio of 6.0. The detected peaks are considered similar across spectra when the m/z values are within a 2000 ppm window.

4. Use hierarchical clustering to classify the spectra, ward algorithm for agglomeration, and a dissimilarity matrix based on the Jaccard distance. The Jaccard index measures similarity between boolean sample sets. The Jaccard distance, which measures dissimilarity between sample sets, is complementary

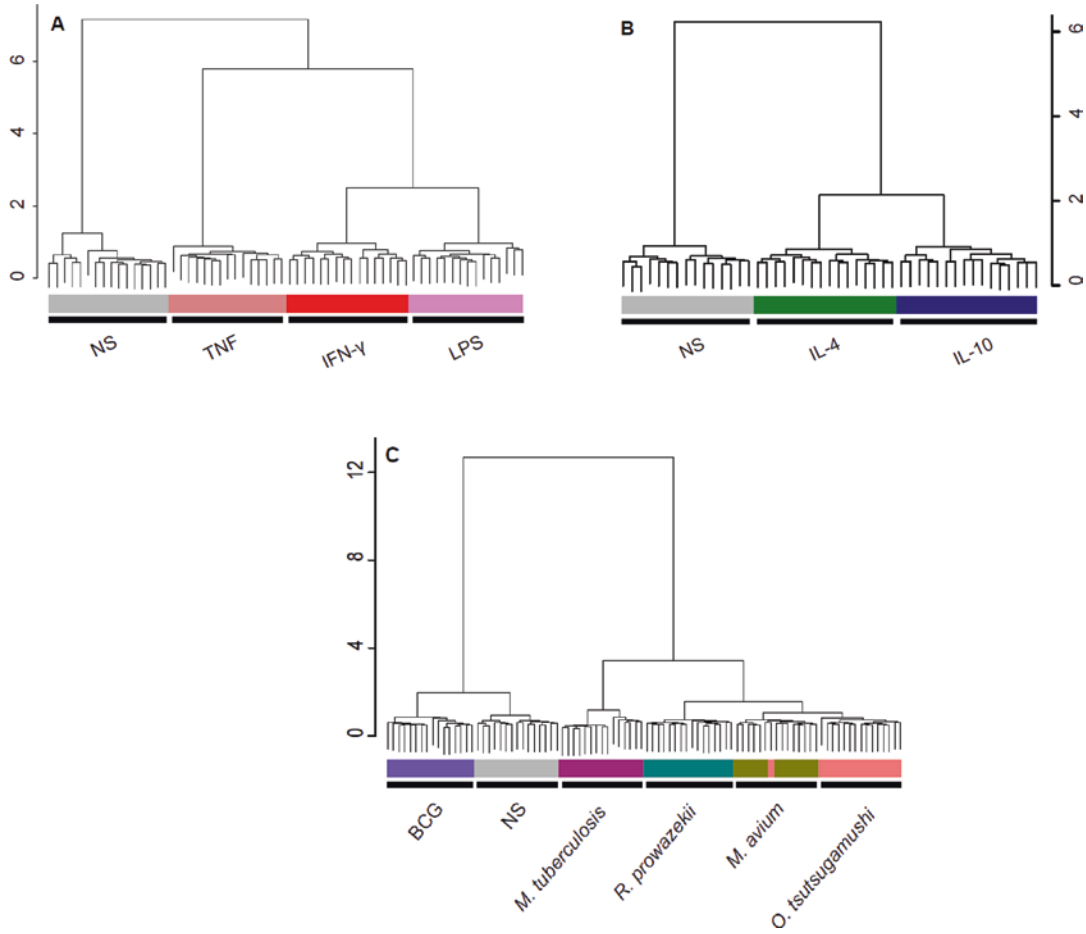


Fig. 5 Hierarchical clustering of activated macrophages. Monocyte-derived macrophages were stimulated with M1-related agonists (**a**), M2-related agonists (**b**), and intracellular bacteria (**c**) for 18 h. Unstimulated macrophages are presented in gray (NS). Figures are reproduced with minor changes from Ouedraogo R, Flaudrops C, Ben Amara A, Capo C, Raoult D, and Mege J-L. Global analysis of circulating immune cells by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. PLoS ONE. 2010;5(10):e13691, under the creative commons license. PLOS applies the Creative Commons Attribution (CC BY) license to articles and other works we publish. If you submit your paper for publication by PLOS, you agree to have the CC BY license applied to your work. Under this Open Access license, you as the author agree that anyone can reuse your article in whole or part for any purpose, for free, even for commercial purposes. Anyone may copy, distribute, or reuse the content as long as the author and original source are properly cited. This facilitates freedom in reuse and also ensures that PLOS content can be mined without barriers for the needs of research

to the Jaccard index and is obtained by subtracting the Jaccard coefficient from 1 or by dividing the difference of the sizes of the union and the intersection of two sets by the size of the union. This procedure was used to discriminate unstimulated (in gray) and stimulated macrophages (in colors) (Fig. 5). Note that the responses of macrophages to different M1 agonists clustered but were not superimposable (Fig. 5a). Similarly, the responses of macrophages to IL-4 and IL-10, two M2 agonists, clustered compared to unstimulated macrophages but were distinct (Fig. 5b). Different intracellular bacteria induced specific signatures (Fig. 5c). Taken together, these results show that MALDI-TOF MS detected the multifaceted activation of macrophages.

4 Notes

1. PMNs must be isolated from remaining RBCs after dextran T500 sedimentation. Lyse RBCs by a 30 s hypotonic shock to obtain pure PMNs. In the absence of lysis, RBC signals were detected in MALDI-TOF MS and masked the detection of PMNs.
2. The stimulation of human MDMs is usually performed in RPMI 1640 supplemented with 10% FCS, 100 UI/mL penicillin, and 50 μg/mL streptomycin.
3. Wash bacteria with PBS to remove the components contained in growth media (such as serum proteins) that may interfere with MALDI-TOF MS spectra.
4. A matrix solution containing crystals does not allow a good ionization of sample molecules and may affect the quality of spectra.
5. A target improperly cleaned may bias the results. It is therefore very important to take the time to carefully clean targets.
6. Cells may agglutinate in the presence of Ca^{2+} and Mg^{2+}. In addition, salts may interfere with MALDI-TOF MS.
7. Rapid and vigorous thawing alters samples, thus affecting the MALDI-TOF MS analysis.
8. Homogeneous deposits are necessary to obtain reproducible and high-quality spectra.
9. Mixing spots with pipettes alters spectrum quality. It is therefore important to respect the proportions and indicated details.
10. Each laboratory needs to construct its own databases before comparing cell populations and looking at unknown samples. We may export our databases to other laboratories.

Acknowledgments

We thank Nicolas Armstrong, Carine Couderc, Philippe Decloquement, and Christophe Flaudrops for their technical assistance.

References

1. Sauget M, Valot B, Bertrand X, Hocquet D (2017) Can MALDI-TOF mass spectrometry reasonably type bacteria? Trends Microbiol 25:447–455
2. Seng P, Rolain J-M, Fournier PE et al (2010) MALDI-TOF-mass spectrometry applications in clinical microbiology. Future Microbiol 5:1733–1754
3. Armengaud J (2017) Defining diagnostic biomarkers using shotgun proteomics and MALDI-TOF mass spectrometry. Methods Mol Biol 1616:107–120
4. Yang H, Chan AL, LaVallo V, Cheng Q (2016) Quantitation of alpha-glucosidase activity using fluorinated carbohydrate array and MALDI-TOF-MS. ACS Appl Mater Interfaces 8:2872–2878
5. Ouedraogo R, Flaudrops C, Ben Amara A et al (2010) Global analysis of circulating immune cells by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. PLoS One 5:e13691
6. Benoit M, Desnues B, Mege J-L (2008) Macrophage polarization in bacterial infections. J Immunol 181:3733–3739
7. Ouedraogo R, Daumas A, Ghigo E et al (2012) Whole-cell MALDI-TOF MS: a new tool to assess the multifaceted activation of macrophages. J Proteome 75:5523–5532
8. Meghari S, Bechah Y, Capo C et al (2008) Persistent Coxiella burnetii infection in mice overexpressing IL-10: an efficient model for chronic Q fever pathogenesis. PLoS Pathog 4:e23
9. Tantibhedhyangkul W, Prachason T, Waywa D et al (2011) Orientia tsutsugamushi stimulates an original gene expression program in monocytes: relationship with gene expression in patients with scrub typhus. PLoS Negl Trop Dis 5:e1028
10. Ryan CG, Clayton E, Griffin WL et al (1988) SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. Nucl Instr Meth Phys Res Sect B 34:396–402



Chapter 24

A Molecular Immunoproteomics Approach to Assess the Viral Antigenicity of Influenza

Kevin M. Downard

Abstract

New surveillance methods employing mass spectrometry (MS) have been developed to characterize the influenza virus and, by extension, other biopathogens at the molecular level. The structure and antigenicity of protein antigens on the surface of the viral capsid are screened in a single step employing the immunoproteomics MS-based approach. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) coupled to gel electrophoresis is used both to identify viral antigens and screen their antigenicity. Evidence that antigen-antibody complexes, and protein complexes more generally, can survive on conventional MALDI targets has allowed both the primary structure and antigenicity of viral strains to be rapidly screened and protein epitopes to be identified with molecular precision. The approach should aid in future screening of the virus and assist in the development of immunogenic peptide constructs as alternative treatments to vaccination over the whole inactivated virus. The assay adds to the repertoire of mass spectrometric approaches for examining antigen–antibody interactions, in particular, and protein complexes, in general, without the need to immobilize, tag, or recover either component.

Key words Influenza, Virus, Surveillance, Antigenicity, Mass spectrometry, Immunoproteomics

1 Introduction

The influenza virus is a leading cause of death linked to the loss of up to 625,000 lives every year [1–3]. It further inflicts illness and suffering the world over, and is responsible for major productive losses, economic and healthcare burdens. Despite a global surveillance strategy [3], the widespread availability and administration of vaccines against the virus [4, 5] and the development of a new generation of antiviral drugs [6], rates of infection continue unabated. Future pandemics, as exemplified by the 2009 swine flu pandemic [7], are a very real threat, and countries have mobilized efforts to stockpile treatments and prepare for outbreaks [8, 9]. New molecular-based surveillance methods employing mass spectrometry (MS) have been developed in this laboratory [10–17] to

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_24, © Springer Science+Business Media, LLC, part of Springer Nature 2019

address limitations of some of the existing approaches employed and help better respond to the virus.

Humans are primarily infected with type A and B influenza. Influenza A viruses also infect birds and other mammals, while type B influenza is almost exclusively a human pathogen. Inactivated forms of common circulating strains of these types form the basis of current vaccines administered against the virus. Viral strains are further subtyped [18, 19] according to the nature of the two surface protein antigens hemagglutinin (H) and neuraminidase (N). The role of the hemagglutinin antigen is to bind the virus particle to sialic acid receptors on the surface of the host cell, while the neuraminidase enzyme releases new budding virus particles that emerge from the host cells and which promote further infection.

The first line of defense offered by the immune response is afforded by antibodies [20]. These are secreted from naive or memory B cells of the host that bind to and inactivate the viral antigens. Alterations in the sequence and structure of viral antigens, as a consequence of antigenic drift and shift, help the virus to elude detection and inactivation by the host's immune system (Fig. 1). These changes are associated with errors produced during the replication of the virus, which lead to mutations in the encoding genes, and the trading of genetic material from one strain to another when humans are infected simultaneously with multiple forms. As humans have no immunity to a strain sufficiently diverged from that to which they were exposed, through natural infection or vaccination, such antigenic drifts and shifts can lead to a local epidemic and even a global pandemic [21].

A worldwide influenza surveillance strategy known as the Global Influenza Surveillance and Response System (GISRS) has been overseen by the World Health Organization for over 65 years since 1952 [3]. The primary goal of this international initiative is to identify and characterize emerging strains of the virus in humans and some animals and recommend the composition of the annual vaccine in the northern and southern hemispheres [3]. Effective

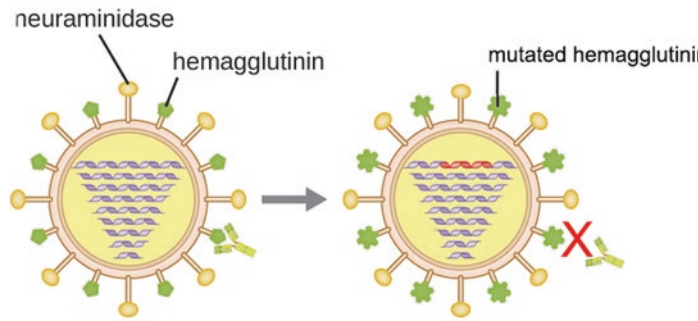


Fig. 1 Antigenic drift introduces mutations in viral proteins (e.g., hemagglutinin) that alters the antigenicity of the virus

vaccines rely on a close match between the current vaccine composition and the antigenicity of circulating strains. As little as a single point mutation in the gene sequence encoding hemagglutinin can result in an existing vaccine being ineffective [22]. The surveillance of the virus in some animals and birds is designed to complement the human studies and aid in the understanding of the ecology of the influenza strains of relevance to human health. The WHO GISRS laboratories screen more than 100,000 specimens positive for influenza per month during the influenza season. Several thousands are submitted to the WHO collaborating centers for antigenic and genetic analysis [3].

The hemagglutination inhibition (HI) assay is the most common assay employed to screen viral isolates [23]. Viruses are first grown in the allantoic fluid of chicken eggs or in cell culture prior to analysis. The inhibition of hemagglutination employs a standardized quantity of the hemagglutinin antigen. The virus is mixed with a suspension of red blood cells, usually isolated from fowl, in a solution of antisera raised to reference strains. The onset of hemagglutination inhibition is identified by the buttoning of red blood cells in the base of a well of the microtiter plate. This results from the attachment of antibody to the hemagglutinin protein that, in turn, prevents the virus from binding with receptors of the erythrocytes. However, it has been found that, particularly for type A H3N2 viruses, successive mutations of Gln226 result in a progressive loss of the ability of the HA to bind chicken erythrocytes [24]. The HI assay also provides no molecular detail as to the nature of viral antigenicity. Furthermore, as the assay is usually interpreted manually (by eye), this contributes to read errors and irreproducibility issues when strains are screened by different individuals in different surveillance laboratories.

Changes to the sequence of the surface antigens are probed employing PCR-based genetic approaches. Since the virus genome is single-stranded RNA, a DNA copy (cDNA) must be prepared using a reverse transcriptase prior to PCR amplification and hybridization sequencing experiments. Microarrays of oligonucleotide probes [25] to known subtypes of the hemagglutinin and neuraminidase antigens have recently been developed to screen viral isolates. Where a new or sufficiently diverged strain is identified, reverse transcriptase-polymerase chain reaction (RT-PCR) sequencing is performed [26]. A primer specific for the hemagglutinin (HA) gene is annealed to the denatured RNA template and extended with reverse transcriptase to synthesize cDNA. After cDNA synthesis, an aliquot of the RT reaction is used for the PCR sequencing cycle. Separation of the PCR products enables the gene sequence to be read [27]. Although automated, genotyping of the influenza virus takes some hours to days to perform. It fails if the primers don't anneal to the desired target sequence that may mutate during the course of the evolution of the virus [28]. The WHO has also identi-

fied problems stemming from the great variation in the precise protocols, extraction kits, and other methodologies in use at National Influenza Centres. These included inconsistent technical performance, inappropriate use of positive controls, laboratory contamination, misinterpretation of results, mismatching of primers and probes, and transcriptional errors.

A mass spectrometry-based approach has been employed which overcomes some limitations of PCR methods for typing and subtyping of the virus [11]. Translated gene sequences for all proteins from a particular type, subtype, and host of the influenza virus sourced from GISAID (Global Initiative on Sharing All Influenza Data) database [29] or a comparable database are aligned. Consensus sequences are established and the frequency of occurrence (Po) of each potential signature peptide is determined using the FluAlign [30] algorithm upon the conservation of each residue of the peptide. The uniqueness of the theoretical monoisotopic mass of each signature peptide is then established using the FluGest algorithm [30] through a comparison with the masses of all proteolytic peptides across all proteins and strains of the virus. The detection of such signature peptides (one or several) within whole-virus or recovered antigen digests enables the strain under investigation to be rapidly and directly typed and subtyped [31–33], their lineage determined [34], and pandemic reasserted viruses distinguished from seasonal strains [35]. The program FluTyper [36] has been developed to automate the process where the mass data is input into an online interface, while the FluShuffle and FluResort algorithms [37] can be used to identify reassorted strains that pose the greatest pandemic risk.

These same mass map datasets can also be used to establish the antigenicity of viral strains where the spectrum of a control sample (untreated, no antibody) is compared with that of an antibody-treated sample. The use of mass spectrometry to screen both the primary structure *and* antigenicity of the virus in a single step [10] in terms of the protein antigens to achieve this was first reported in 1999 [15]. This immunoproteomics approach [15], originally applied to whole virus [12], was later advanced for gel-purified antigens [13, 14] in order to improve sequence coverage and the likelihood of epitope identification (Fig. 2). Important to the success of the approach is the realization that peptides representing an antigenic determinant can remain bound to antibody throughout immunological treatment, MALDI sample deposition and ablation [10, 12–15], and mass spectral analysis providing certain experimental conditions are implemented [10, 12, 17]. The antigenicity of component antigens can then be assessed by a comparison of a pair of their MALDI mass spectra [10, 12–15] before and after their treatment with monoclonal antibodies. The methods involved in this immunoproteomics approach [10, 12–15] are described herein both to highlight the advantages of the method and to outline the experimental protocols.

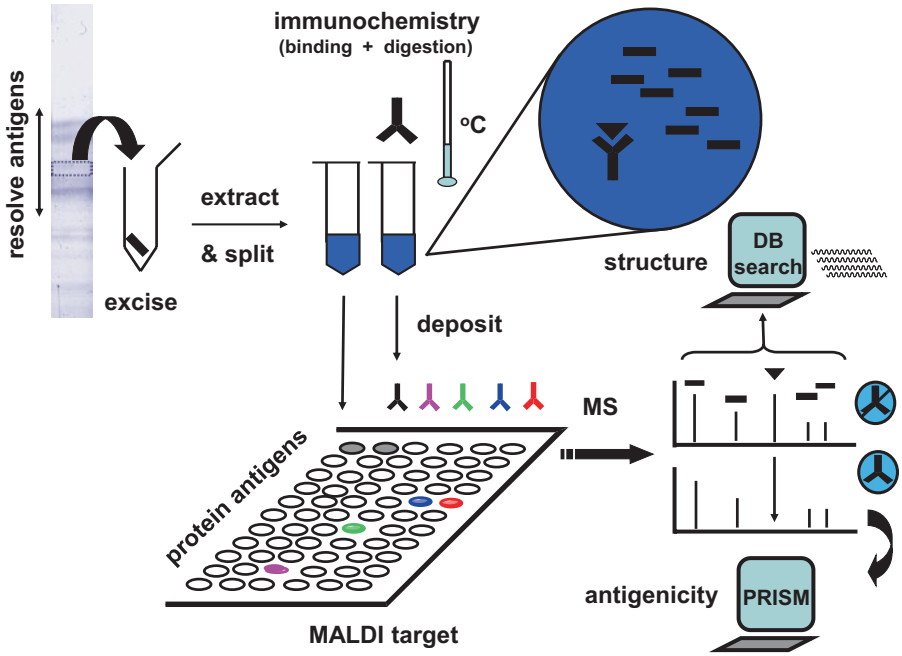


Fig. 2 Schematic representation of the immunoproteomics approach to survey both the primary structure and antigenicity of antigens of the influenza virus. Gel-resolved antigens are treated with monoclonal antibodies prior to or following proteolytic digestion. The products are analyzed directly by using MALDI mass spectrometry. Antigens are identified through searches of general or specialized flu databases [29] with the mass fingerprint data. The PRISM algorithm [24] aids in the identification of epitopic peptides that bind antibody

2 Materials

2.1 Viral Isolates

Inactivated viruses were obtained from commercial sources or the World Health Organization Collaborating Centre for Reference and Research on Influenza, Melbourne. Viruses were inactivated with 0.005% merthiolate and purified by ultracentrifugation using a 10–40% sucrose gradient.

2.2 Monoclonal Antibodies

Monoclonal antibodies obtained commercially were derived from the hybridization of myeloma and spleen cells of mice immunized with the relevant serotype, prepared in a miniPERM bioreactor in serum-free medium, and/or obtained from commercial sources.

2.3 Gel Electrophoresis

1. Acrylamide.
2. Ammonium persulfate.
3. Gel buffer (25 mM Tris–HCl, 192 mM glycine at pH 8.3 with 0.1% (w/v) of sodium dodecyl sulfate for denaturing gels).
4. Mini gel box apparatus and power supply.

5. Gel stain solution [0.25% (w/v) Coomassie blue (R-250) solution in water (53% v/v), methanol (40% v/v), and acetic acid (7% v/v)].
6. Destain solutions [acetonitrile and water (50% v/v) and acetonitrile and 10–50 mM ammonium bicarbonate (50% v/v)].
7. Digital gel imager.

2.4 Enzymatic Digestion and Peptide Recovery

1. Ammonium bicarbonate buffer (50 mM, pH 8).
2. Sequence-grade protease (e.g., trypsin).
3. PNGase F.
4. Trifluoroacetic acid.
5. Bath sonicator.
6. Centrifugal vacuum concentrator.

2.5 Immunochemistry

1. PBS buffer (20 mM phosphate, 100 mM NaCl, pH 7.8).

2.6 Mass Spectrometry

1. MALDI-based mass spectrometer.
2. Matrix solution [saturated solution of α-cyano-4-hydroxycinnamic acid in acetonitrile and water (70:30 v/v)].

2.7 Data Analysis

1. GISAID: Global Initiative on Sharing All Influenza Data [29], FluServer App https://www.gisaid.org/.
2. PRISM (Protein Interactions from the Spectra of Masses) software [38].

3 Methods

3.1 Gel Electrophoresis

The inclusion of a stage of separation of viral antigens [13, 14], over the use of whole virus [12], improves the sequence coverage achieved in the MALDI mass map [39]. The presence of additional antigens and components can result in the overlap or suppression of ion signals for a particular antigen. This impacts on the ability to correctly identify the antigen, establish differences in its sequence from a like antigen of a reference strain, and follow the binding of antigen to antibody by monitoring the reduction in an epitopic peptide signal relative to nonbinding peptides.

In order to identify and characterize discontinuous epitopes, antigen separation is achieved using native gel electrophoresis without the use of sodium dodecyl sulfate. Gradient gels afford better separations for this application. Protein antigen bands are subsequently identified using a sacrificial stain-aligned band in

order to prevent the denaturation of recovered antigen with the stain and during the destaining of bands in organic solvents. Whole antigens are recovered from the native gel using a protocol developed for this purpose that employs bath sonication [40]. Gel plugs are sliced into small pieces, pulverized, and sonicated in a bath sonicator for 3 × 15 min with the samples repeatedly cooled over ice to prevent heating of the solution. Protein recoveries of between 60 and 80% have been achieved for model proteins [40]. The integrity of their structures following this treatment was confirmed by electrospray ionization mass spectrometry (ESI-MS) and circular dichroism (CD) spectroscopy [40].

Solutions of influenza strains (at 1 μg/μl) in gel buffer were loaded in separate lanes of a commercial or cast 8–17.5% polyacrylamide gel and run with gel buffer, in the presence or absence of 0.1% w/v of sodium dodecyl sulfate, for 2–3 h at 150 V. Gels, or a single reference lane of the gel, were stained in stain solution overnight and destained for 2 h. Despite the electrophoretic separation stage, some bands may still contain more than one antigen particularly where native gel electrophoresis is employed. Even in the case of denaturing PAGE, the hemagglutinin and nucleoprotein antigens of type H1N1 strains share similar molecular weights and are found to migrate to similar positions on the gel. Nonetheless, even the incomplete purification of antigens from whole virus improves their sequence coverage in the mass map that subsequently aids in the identification of epitopic domains.

3.2 Immunobinding and Proteolytic Digestion

The bands containing the viral antigens hemagglutinin or neuraminidase are excised, cut into small pieces, and transferred into separate pre-washed (with acetonitrile) tubes for extraction of the protein either before or after digestion (*see* **Note 1**). Whole antigens are recovered by sonication [40] and with the aid of trifluoroacetic acid in the case of in-gel digested protein (*see* **Note 2**). Antigens are treated with monoclonal antibodies prior to or after enzymatic digestion [12–14] subject to the nature of the predicted epitope. Treatment with monoclone precedes limited digestion of the immune complex [41] where the nature of the epitope is unknown. Solutions of monoclonal antibodies in PBS buffer are treated with antigen at a 2:1 antibody-to-antigen mole ratio for a period of 24 h at 4 °C, following antigen recovery from a native gel [40] or after its digestion in gel [13, 14]. After resuspension in ammonium bicarbonate buffer, sequence-grade endoproteinase is added to effect the limited proteolysis of antibody–antigen complexes, or digestion of antigen ahead of antibody treatment, for a period of up to 24 h at 37 °C, where aliquots removed at different time points can be used in a time-course immunoassay to derive binding rate data.

3.3 Mass Spectrometry and MS Data Analysis

A portion of the antibody reaction mixture and untreated control (1 μL) are diluted in a saturated solution of α-cyano-4-hydroxycinnamic acid (3 μL) in acetonitrile and water (70:30 v/v) and deposited onto the sample stage (1 μL per spot) (*see* **Note 3**). The samples are analyzed by MALDI-MS using a 337 nm nitrogen laser operating at approximately 10 μJ (*see* **Note 4**) employing time-delayed ion extraction (delay time ~10 μs) on a linear or reflecting time-of-flight mass analyzer [41] (*see* **Note 5**). The MALDI mass spectra are acquired in seconds affording a significant advantage over the time required for analysis in other analytical and spectroscopic approaches. The resulting spectra, each acquired from an average of some 100 laser shots from across the target in order to ensure the representative nature of the peptides within the sample and also improve ion statistics, are compared manually and/or with the aid of a computer program (PRISM) developed in-house for this purpose [37]. Spectra of the antigen-only sample are compared with *m/z* values for proteolytic peptides predicted for the protein based on translated gene or protein sequences derived from the GISAID or similar database [29].

The identification of protein epitopes or determinants is established based on the selective reduction in the relative area of their ion signals of the spectra recorded for the no-antibody control sample versus the antibody-treated sample. An absolute reduction of at least 10–15% in relative area is required to establish the identity of an antigenic peptide(s) in order to compensate for experimental errors associated with the variation of mass spectra obtained from different samples (*see* **Note 6**). Replicate experiments from common and diverged strains are used to verify the results. The latter also helps establish critical binding residues where differences among antigen sequences within an epitope may impact the ability of the peptide to bind monoclone.

An algorithm [38] has been developed to compare pairs of mass spectra for the no-antibody and antibody-treated sample to aid with the identification of protein epitopes. Known as PRISM, as it identifies PRotein Interactions from the Spectra of Masses, the algorithm features an easy-to-use, intuitive graphical user interface. The data is input as a list of mass or *m/z* values for the peptide ions and their abundances. Each pair of spectra is processed to remove any *m/z* values of ions found in only one spectrum (within a specified *m/z* error; default = 0.1) and that are of low signal-to-noise relative to the most abundant base peak (default 1%). The algorithm then compares changes in the area under each of the ion signals relative to a constant peak. The constant peak is determined by first measuring the ratio of areas of all adjacent peaks $n + 1/n$ (in terms of *m/z*) in both spectra and establishing the ion peak n for which the absolute difference is the smallest. The value for the relative area of the constant peak to the most abundant base peak in the control (no-antibody) spectrum is then used to establish an

imaginary peak in the spectrum of the antibody-treated sample. This is achieved such that the relative area of the constant peak to the base peak in the control spectrum is equal to the ratio of area of the constant peak to the imaginary peak in the spectrum of the antibody-treated sample. This establishes a common reference peak with which to compare relative areas across the two different mass spectra. Areas under the ion signals are computed relative to the base peak in spectrum of the control and relative to the imaginary peak within the spectrum of the antibody-treated sample. Absolute changes in relative area among common ions in both spectra that are greater in value than 10% (default value) to 15% represent possible antibody-binding peptides. The *m/z* values for these ions are output together with the identity of the constant peak to enable the spectra to be independently assessed.

To illustrate a typical set of mass spectra and their analysis, MALDI mass spectral data for the gel-recovered hemagglutinin of H1N1 influenza strain (Beijing 262/95) before and after treatment with a monoclonal antibody raised the antigen of this serotype are shown in Fig. 3 [13]. Note that recovered antigen also contains some nucleoprotein (NP) that was not completely separated by gel electrophoresis. Each spectrum provides a mass

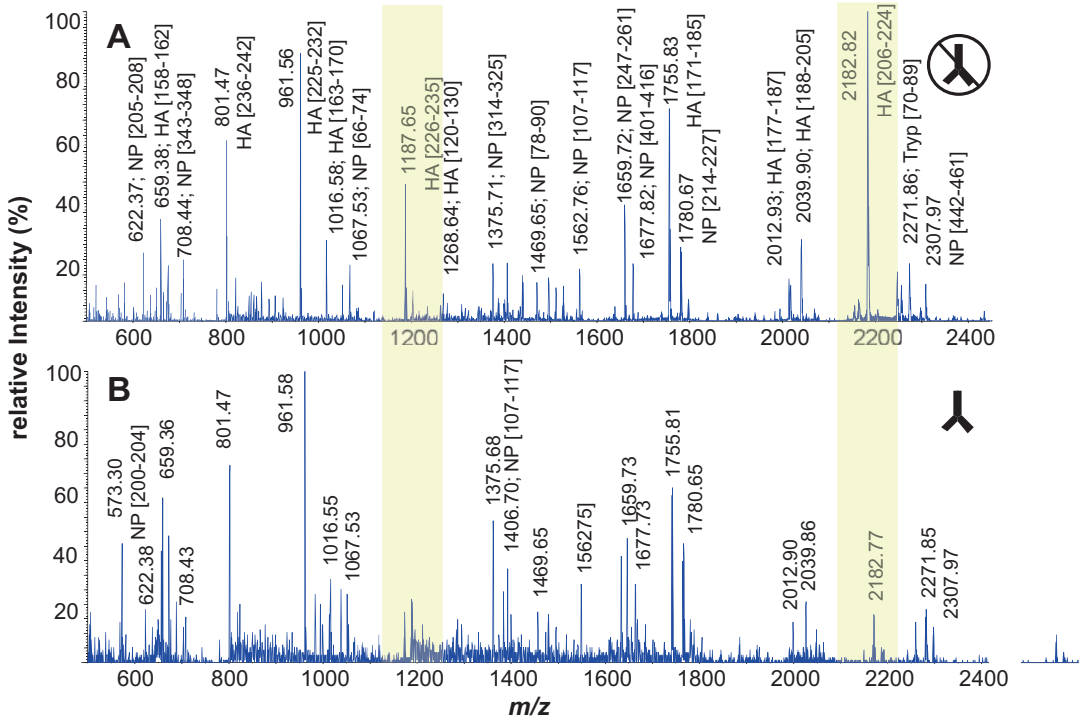


Fig. 3 MALDI mass spectra of gel-recovered hemagglutinin after tryptic digestion of the Beijing 262/95 type A strain (A) without antibody and (B) after 24 h of incubation with a monoclonal antibody raised to a H1N1 serotype [11]

"fingerprint" or map of the viral antigen(s) that can be used in its own right to search protein databases [39] to confirm or identify differences in the sequences of the component antigens. A Mascot peptide mass fingerprint search of the MSDB nonredundant database correctly establishes the identity of the antigens in the band as a mixture of both nucleoprotein (of a diverged Hong Kong H1N1 strain) and the hemagglutinin antigen (Fig. 4). Differences in amino acid sequence can be easily be confirmed by generating tandem (or MS/MS) spectra for any or all other peptides including epitopic peptides [12–15]. On most mass spectrometers, this can be performed in an automated manner where ions of a specified *m/z* value, or those above a signal threshold, are isolated in the first

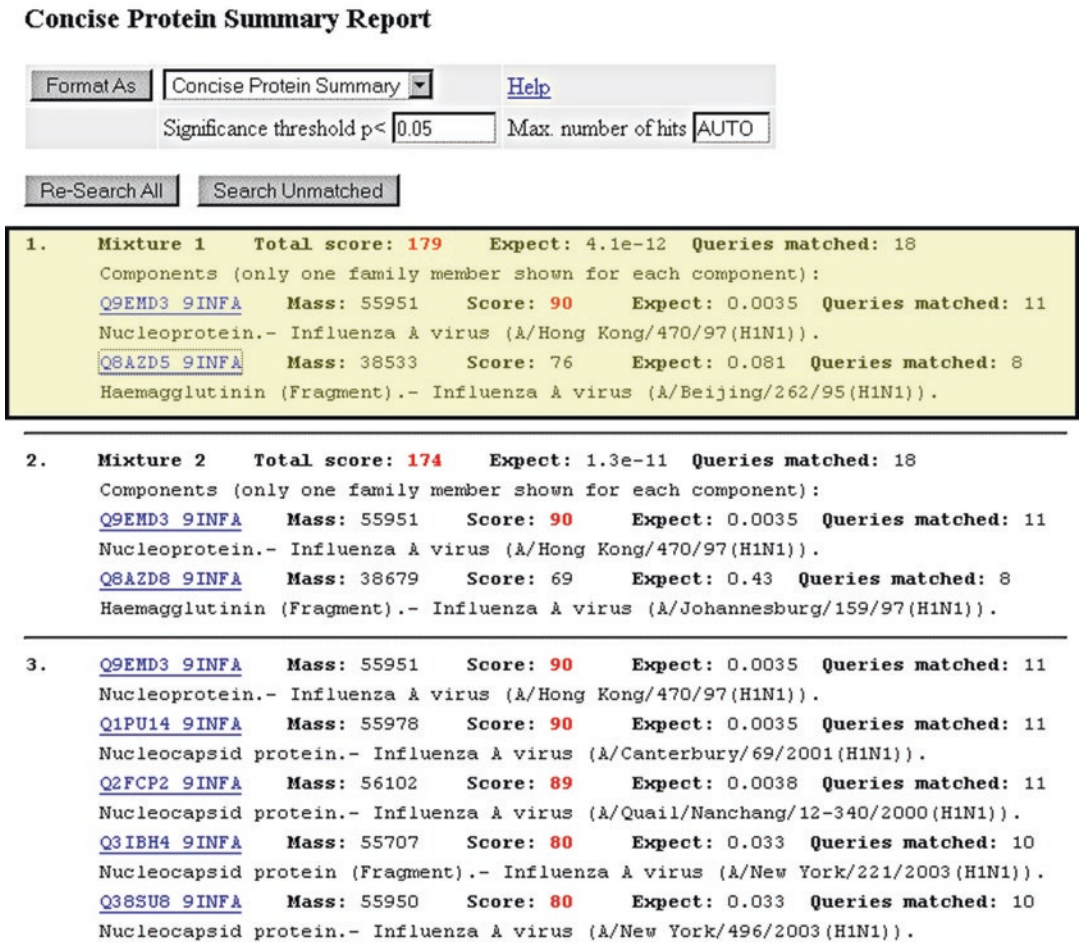


Fig. 4 Mascot peptide mass fingerprint search output for the three highest scored protein candidates obtained from a search of the MSDB database with the mass spectral data from Fig. 3A. The MSDB database is a composite, non-identical protein sequence database compiled from entries of the PIR, TrEMBL, GenBank, SwissProt, and NRL3D databases at Imperial College, London

mass analyzer and subjected to dissociation in MS/MS experiments [42].

Despite an incomplete separation of the antigens by gel electrophoresis, the peptide ion comprising residues 206–224 of hemagglutinin at *m/z* 2182.8 is seen to significantly reduce in its relative area (94% absolute change) in the presence of antibody due to its interaction with it and the preservation of the peptide complex on the MALDI target. All other peptide ion signals remain relatively unchanged within experimental area (some ≤15%) with the exception of the peptide ions at *m/z* 1187.65. This peptide corresponds to residues 226–235 of the hemagglutinin antigen that flanks the C-terminal end of the epitope and also binds to a lesser degree. The location of the epitope is in accord with earlier results obtained for a diverged H1N1 strain using whole-virus digests [12].

The ability of the PRISM algorithm to identify the epitopic peptide from such pairs of mass spectra without human intervention is illustrated in Fig. 5. The two mass spectra shown in Fig. 3 are input into the algorithm as *m/z* values versus peak area data. After noise reduction and the removal of ions not common to both spectra, the algorithm compares the relative area of ions across both spectra using the approach described above. Using a cutoff of 15% (*see* **Note 7**), only three peptide ion signals are found whose relative areas change by a level that exceeds this value. These include the two peptides identified above, validating the performance of the algorithm, and an additional peptide comprising residues at 120–130 of the hemagglutinin antigen (*m/z* 1268.6) which shares sequence homology with the epitopic peptide at *m/z* 2182.82 [13].

3.4 Comparison of Results with HI Assay

The successful application of this immunoproteomics approach for type A H1N1 strains [13] has been extended to other strains of the virus including three diverged H3N2 type A strains [14] and two type B strains [31], and the results directly have been compared with those obtained in a hemagglutination inhibition (HI) assay [43]. The trypsin digested HA1 hemagglutinin subunit of the A/New Caledonia/20/99 strain, separated by SDS-PAGE, is shown in Fig. 6 together with the same mixture treated with two monoclonal antibodies mAb128 (anti-influenza A H1N1) and mAb154 (anti-influenza B). Despite the band also containing a proportion of nucleoprotein (NP), the selective binding of the HA peptide at *m/z* 961.5 comprising residues 225–232 is evident by an absolute decrease in the relative area of its ion signal of 36% compared with that for the no-antibody and mAb154-treated samples, respectively. The results of the HI assay show that mAb128 binds hemagglutinin of the A/New Caledonia/20/99 strain at antibody levels of 0.2 μg or above, resulting in a pooling of erythrocytes in the bottom of microtiter plate. The anti-influenza B antibody mAb154 in contrast fails to inhibit hemagglutination in

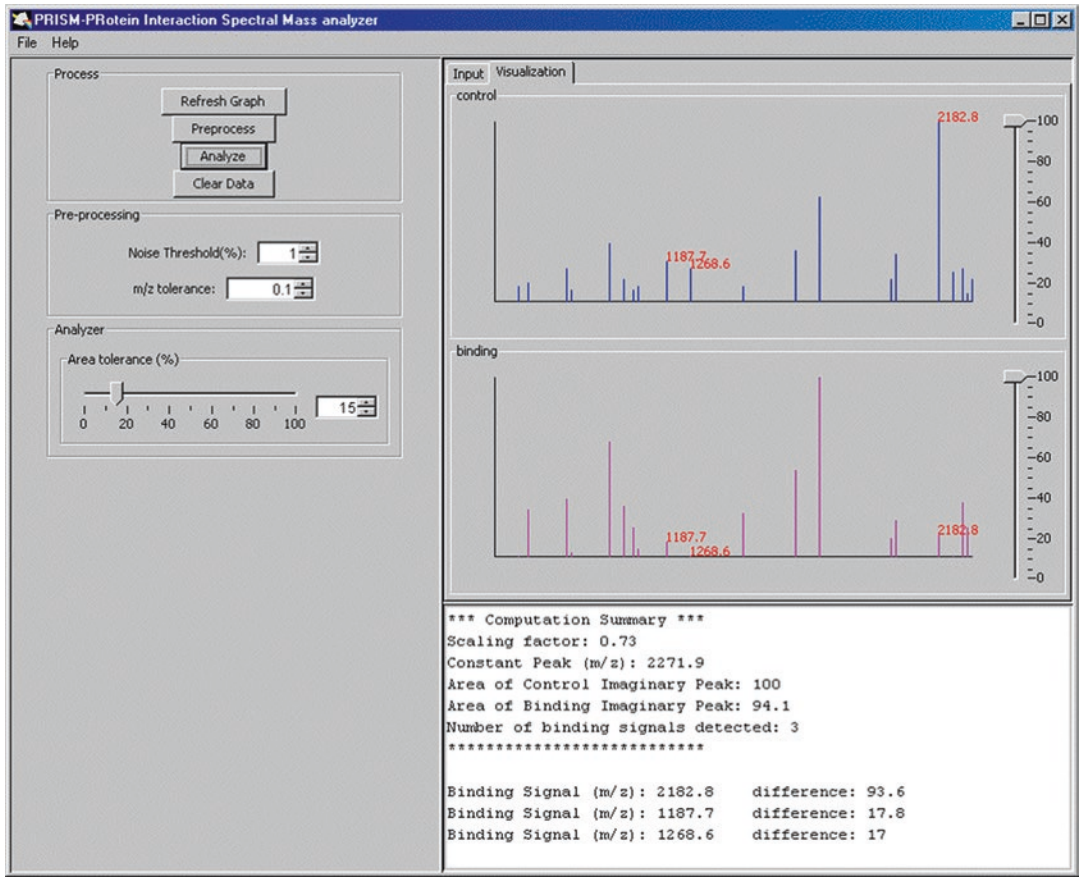


Fig. 5 Graphical user interface (GUI) of the PRISM algorithm [24] showing the output for the mass spectral data shown in Fig. 2

the presence of the type A/New Caledonia/20/99 strain across all antibody concentrations but does so against the type B/Victoria(Vic)/504/2000 strain at levels above 0.04 μg.

3.5 Application to the Determination of Affinity Constants

The implementation of time-course experiments enables the relative rates that epitopic peptides bind with antibodies to be assessed, in both noncompetitive and competitive immunoassays, by measuring the levels of unbound antigenic peptide regions that remain in the sample over time. A time-course MALDI-MS immunoassay has been shown to be able to detect differences in the relative rates of binding of peptides, both from within and across epitopic domains, with antibodies in noncompetitive and competitive experiments [44].

As an example, the gel-separated hemagglutinin of a H3N2 type A/Kiev/301/94 strain was found to bind a monoclone raised to HA in the vicinity of residues 158–166 (m/z 1000.5) and 109–125 (m/z 2067.9) due to their progressive decrease in relative intensity following increased time exposures. The rates of binding

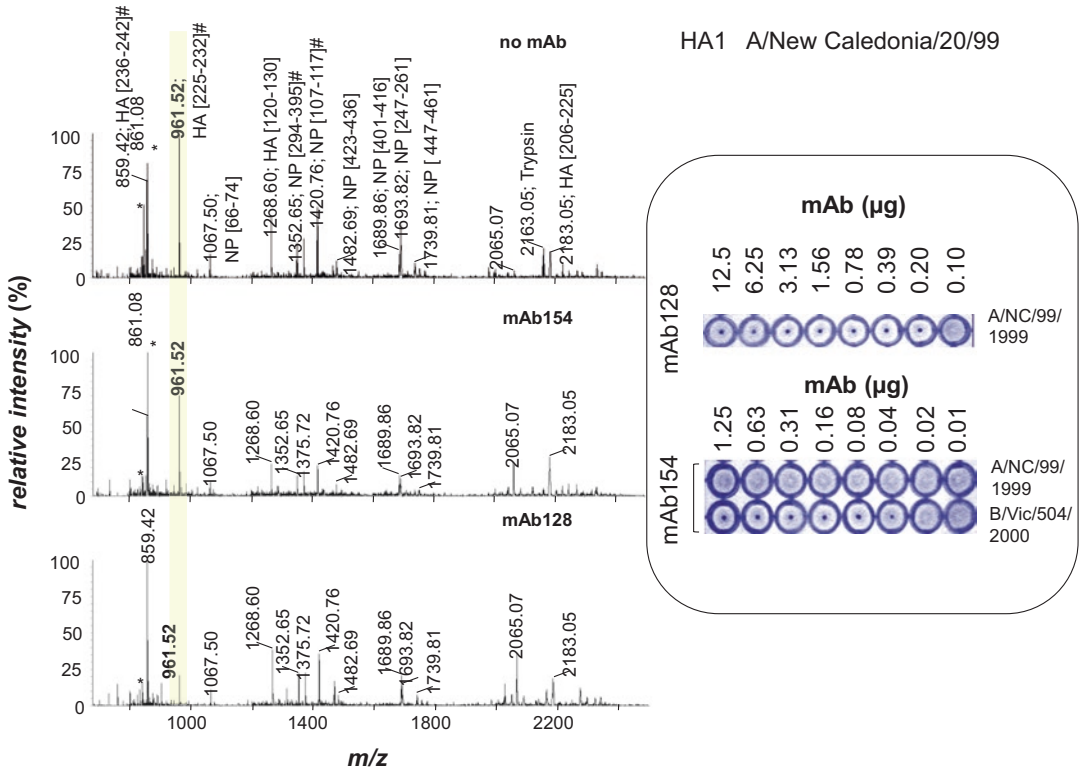


Fig. 6 MALDI mass spectra of the tryptic peptides released from the HA1 band, untreated (control) or treated with an anti-influenza A (mAb 128) and anti-influenza B (mAb154) antibody and companion hemagglutination inhibition assay (boxed). The binding peptide at *m/z* 961.5 is highlighted [43]. Note that the ion signal at *m/z* 861.1 (denoted by an asterisk) is associated with matrix cluster ions that appear with variable intensity due to their sporadic production during the ionization process

were determined from the decay curves (Fig. 7) to be 0.24 and 0.05 h^{-1}, respectively, such that the relative rate of binding in this competitive assay favors the smaller peptide (residues 158–166) by a ratio approximately 5:1 [44]. The smaller peptide of nine residues is consistent with that typically implicated as constituting an IgG antibody-binding domain. The preferential binding of this peptide may also be afforded by its smaller size, which reduces unfavorable less steric restrictions. This is supported by the observation that a shorter segment of residues 119–125 (of sequence PDYASLR) comprising the C-terminus of the larger binding peptide was also found to bind antibody. To identify critical binding residues in this segment, eight peptides were synthesized that successively replaced the aspartic acid, tyrosine, serine, and arginine residues. Each residue was replaced with a more hydrophobic residue, but one with a side chain of a comparable size. Thus, tyrosine is substituted with a phenylalanine residue, the aspartic acid and serine residues were substituted successively with alanine residues, and the arginine residue was substituted successively with a leucine

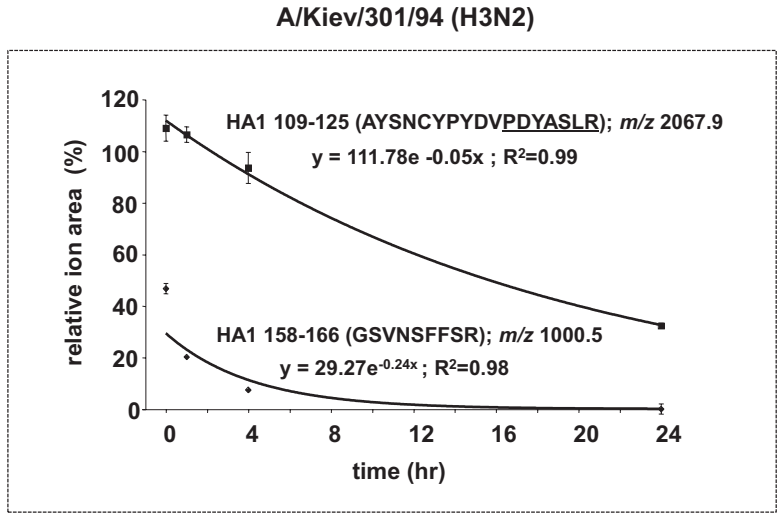


Fig. 7 Time-course MALDI mass spectrometry immunoassay showing the decrease in a relative area of ions for two epitopic peptides comprising residues 109–125 and 158–166 of the hemagglutinin HA1 subunit of the A/Kiev/301/94 (H3N2) strain after binding with a monoclonal antibody over 24 h [44]. Data points represent the average values from duplicate experiments with three separate MS spectral datasets recorded per experiment

residue or a glutamic acid residue. The latter changes the charge character of the C-terminal residue from as highly basic to an acidic residue. The synthesized peptides were individually assessed to bind antibody based on a measure of their relative area after treatment in the presence of two control peptides. Four peptides, including the unsubstituted epitopic peptide, were found to bind the antibody, while the other four peptides did not. The three peptides with sequences PDYASLE (where arginine is replaced by glutamic acid), PDYASLL (where arginine is replaced by leucine), and PDYAALR (where serine is replaced by alanine) were found to decrease in relative area by 89%, 47%, and 68%, respectively. Noncompetitive time-course immunoassays for each of the four binding peptides allowed rate coefficients to be measured that ranged from 0.52 to 1.0 min^{-1}, representing a maximum two-fold difference. When each of the three peptides was examined in competition with the unsubstituted peptide, the differences were not significant within experimental error.

3.6 Conclusion

This MS-based immunoassay adds to the repertoire of mass spectrometric approaches for examining antigen–antibody interactions, in particular, and protein complexes, in general, without the need to immobilize, tag, or recover either component. It has particular utility, as illustrated in this review, to improve the surveillance of the influenza given the continued impact of the virus on human health [45].

4 Notes

1. Endoproteinases should be chosen with care in order to effect site-specific limited digestion of the antigen–antibody complex or antigen alone so as to refine epitopic domains without cleaving epitopic peptides. It has been shown in one study [9] that where digestion was performed with chymotrypsin, a linear epitope (identified following a tryptic digestion) within H1 hemagglutinin was cleaved and no binding peptides were detected. If no epitopic peptides are detected in one experiment, a second protease should be chosen to assess this issue.
2. Proteolytic digestion should aim to achieve the maximum antigen sequence coverage in order to follow changes in the relative area of ion signals for peptides across the entire protein. The use of two experiments with two different endoproteinases can assist with this goal. In practice, 100% coverage is rarely achieved even for highly purified protein standards in mass mapping experiments [25].
3. MALDI matrix additives such as acetic acid and trifluoroacetic acid must be avoided since these have been shown to induce dissociation of peptide–antibody complexes leading to the detection of epitopic peptides in the antibody-treated sample [9].
4. Laser powers should be kept as low as possible in order not to disrupt peptide–antibody complexes on the MALDI target. The higher extraction potentials required to directly detect immune complexes [16] are not required where such complexes are preserved on the target surface.
5. Any MALDI-based mass spectrometer from TOF, hybrid-TOF, or other mass analyzers (e.g., FT-ICR) can be employed in this approach where appropriate consideration is given to analyzer scan rates and their dynamic mass range.
6. A 10–15% absolute change in the relative area of ion signals has been demonstrated to be an appropriate cutoff value to efforts to detect binding peptides in order to allow for experimental fluctuations that cause slight ion intensity variations in mass spectral data for different samples [9–12].
7. The spectra acquired should be representative of the total peptide composition on the MALDI target. In this respect, the plated sample should be as homogeneous as possible, and/or sample should be ablated from across the surface by moving the laser position or target.

References

1. Nicholson KG, Wood JM, Zambon M (2003) Influenza. Lancet 362:1733–1745
2. Barry JM (2004) The great influenza: the epic story of the deadliest plague in history. Penguin Viking, New York, NY
3. World Health Organization Global Influenza Surveillance Network. http://www.who.int/influenza/surveillance_monitoring/en/
4. Stephenson I, Nicholson KG (2001) Influenza: vaccination and treatment. Eur Respir J 17:1282–1293
5. Girard MP, Cherian T, Pervikov Y, Kieny MP (2005) A review of vaccine research and development: human acute respiratory infections. Vaccine 23:5708–5724
6. Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, Garten RJ, Gubareva LV, Xu X, Bridges CB, Uyeki TM (2009) Emergence of a novel swine-origin influenza a (H1N1) virus in humans. N Engl J Med 360:2605–2615
7. Garman E, Laver G (2004) Controlling influenza by inhibiting the virus's neuraminidase. Curr Drug Targets 5:119–136
8. Webster RG, Walker EJ (2003) Influenza: the World is teetering on the edge of a pandemic that could kill a large fraction of the human population. Am Sci 91:122
9. Horimoto T, Kawaoka Y (2005) Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Microbiol 3:591–600
10. Downard KM, Morrissey B, Schwahn AB (2009) Mass spectrometry analysis of the influenza virus. Mass Spectrom Rev 28:35–49
11. Downard KM (2013) Proteotyping for the rapid identification of pandemic influenza virus and other biopathogens. Chem Soc Rev 42:8584–8595
12. Kiselar JG, Downard KM (1999) Antigenic surveillance of the influenza virus by mass spectrometry. Biochemistry 38:14185–14191
13. Morrissey B, Downard KM (2006) A proteomics approach to survey the antigenicity of the influenza virus by mass spectrometry. Proteomics 6:2034–2041
14. Morrissey B, Streamer M, Downard KM (2007) Antigenic characterisation of H3N2 subtypes of the influenza virus by mass spectrometry. J Virol Methods 145:106–114
15. Kiselar JG, Downard KM (1999) Direct identification of protein epitopes by mass spectrometry without immobilization of antibody and isolation of antibody-peptide complexes. Anal Chem 71:1792–1799
16. Kiselar JG, Downard KM (2000) Preservation and detection of specific antibody-peptide complexes by matrix-assisted laser desorption ionization mass spectrometry. J Am Soc Mass Spectrom 11:746–750
17. Downard KM (2016) Indirect study of non-covalent protein complexes by MALDI mass spectrometry. Origins, advantages and applications of the "intensity-fading" approach. Mass Spectrom Rev 35:559–573
18. Zhang WD, Evans DH (1991) Detection and identification of human influenza viruses by the polymerase chain reaction. J Virol Methods 33:165–189
19. Wright KE, Wilson GA, Novosad D, Dimock C, Tan D, Weber JM (1995) Typing and subtyping of influenza viruses in clinical samples by PCR. J Clin Microbiol 133:1180–1184
20. Ruben FL, Jackson GG, Gotoff SP (1973) Humoral and cellular response in humans after immunization with influenza vaccine. Infect Immun 7:594–596
21. Wood JM (2002) Selection of influenza vaccine strains and developing pandemic vaccines. Vaccine 20:B40–B44
22. Kodihalli S, Justewicz DM, Gubareva LV, Webster RG (1995) Selection of a single amino acid substitution in the hemagglutinin molecule by chicken eggs can render influenza a virus (H3) candidate vaccine ineffective. J Virol 69:4888–4897
23. Julkunen I, Pyhala R, Hovi T (1985) Enzyme immunoassay, complement fixation and hemagglutination inhibition tests in the diagnosis of influenza a and B virus infections. Purified hemagglutinin in subtype-specific diagnosis. J Virol Methods 10:75–84
24. Medeiros R, Escriou N, Naffakh N, Manuguerra JC, van der Werf S (2001) Hemagglutinin residues of recent human a(H3N2) influenza viruses that contribute to the inability to agglutinate chicken erythrocytes. Virology 289:74–85
25. Li J, Chen S, Evans DH (2001) Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. J Clin Microbiol 39:696–704
26. O'Connell J (ed) (2002) RT-PCR protocols, Methods in Molecular Biology, vol 193. Humana Press, NJ
27. Poddar SK (2002) Influenza virus types and subtypes detection by single step single tube multiplex reverse transcription-polymerase chain reaction (RT-PCR) and agarose gel electrophoresis. J Virol Methods 99:63–70

28. Yang J-R, Kuo C-Y, Huang H-Y, Wu F-T, Huang Y-L, Cheng C-Y, Su Y-T, Chang F-Y, Wu H-S, Liu M-S (2014) Newly emerging mutations in the matrix genes of the human influenza a(H1N1)pdm09 and a(H3N2) viruses reduce the detection sensitivity of real-time reverse transcription-PCR. J Clin Microbiol 52:76–82
29. Shu Y, McCauley JM (2017) GISAID: global initiative on sharing all influenza data – from vision to reality. Euro Surveill 22:30494
30. Schwahn AB, Wong JWH, Downard KM (2009) Subtyping of the influenza virus by high resolution mass spectrometry. Anal Chem 81:3500–3506
31. Schwahn AB, Wong JWH, Downard KM (2009) Signature peptides of influenza nucleoprotein for the typing and subtyping of the virus by high resolution mass spectrometry. Analyst 134:2253–2261
32. Schwahn AB, Wong JWH, Downard KM (2010) Typing of human and animal strains of influenza virus with conserved signature peptides of matrix M1 protein by high resolution mass spectrometry. J Virol Methods 165:178–185
33. Nguyen AP, Downard KM (2013) Subtyping of influenza neuraminidase using mass spectrometry. Analyst 138:1787–1793
34. Schwahn AB, Downard KM (2011) Proteotyping to establish the lineage of type a H1N1 and type B human influenza virus. J Virol Methods 171:117–122
35. Schwahn AB, Wong JWH, Downard KM (2010) Rapid differentiation of seasonal and pandemic H1N1 influenza through proteotyping of viral neuraminidase with mass spectrometry. Anal Chem 82:4584–4590
36. Wong JWH, Schwahn AB, Downard KM (2010) FluTyper-an algorithm for automated typing and subtyping of the influenza virus from high resolution mass spectral data. BMC Bioinformatics 11:266
37. Lun ATL, Wong JWH, Downard KM (2012) FluShuffle and FluResort-new algorithms to identify Reassorted strains of the influenza virus by mass spectrometry. BMC Bioinformatics 13:208
38. Ho JWK, Morrissey B, Downard KM (2007) A computer algorithm for the identification of protein interactions from the spectra of masses (PRISM). J Am Soc Mass Spectrom 18:563–566
39. Pappin DJC, Hojrup P, Bleasby AJ (1993) Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol 3:327–332
40. Mackun K, Downard KM (2003) Strategy for identifying protein-protein interactions of gel-separated proteins and complexes by mass spectrometry. Anal Biochem 318:60–70
41. Sheshberadaran H, Payne LG (1988) Protein antigen– monoclonal antibody contact sites investigated by limited proteolysis of monoclonal antibody-bound antigen: protein "footprinting". Proc Natl Acad Sci 85:1–5
42. Downard KM (2004) Tandem mass spectrometry. In: Mass spectrometry – a foundation course, chapter 4. Royal Society of Chemistry, Cambridge, UK
43. Schwahn AB, Downard KM (2009) Antigenicity of a type a influenza virus through a comparison of hemagglutination inhibition and mass spectrometry immunoassays. J Immunoassay Immunochem 30:245–261
44. Morrissey B, Downard KM (2008) Kinetics of antigen-antibody interactions employing a MALDI mass spectrometry immunoassay. Anal Chem 80:7720–7726
45. Downard KM, Morrissey B (2007) Fingerprinting a killer-surveillance of the influenza virus by mass spectrometry. Analyst 132:611–614



Chapter 25

Immunoproteomic Biomarkers: From Publication to Personalized Medicine

Devanand M. Pinto

Abstract

As the number of immunotherapies increases, so does the need for biomarkers that can aid in identifying an optimal therapy or combination therapy for patients. These predictive biomarkers are of enormous value to patients but present unique challenges to researchers due to the complexity of the immune system and the variability of individual patient molecular profiles. This chapter draws on recent examples of the use of biomarkers to explore the range of phenotypes encountered in immunotherapy trials for the treatment of neoplastic disease. These examples are discussed in the context of immunoproteomic analysis with a particular focus on the unique challenges that are presented when a high dimensionality technique such as immunoproteomics is applied to study a complex system, the immune system in this case. In order to overcome these challenges, immunoproteomic researchers must pay close attention to study design in order to ensure that the results are not only statistically valid but also that the biomarker strategy as a whole is compatible with the standard of care. We propose that, in spite of its limitations, the use of immunoproteomic analysis of liquid biopsies may present a unique opportunity for translation of immunoproteomic biomarkers to the clinic.

Key words Biomarker validation, Evidence-based medicine, Immunoproteomics, Liquid biopsy

1 Introduction

Validated biomarkers can be used to detect the presence of disease, determine the likely course of disease, or choose a treatment most likely to provide a significant response. These diagnostic, prognostic, or predictive biomarkers each take a unique path from initial discovery to validation depending on their intended use. Biomarker discovery usually begins with proof-of-concept, single-institution studies that include a small number of patients; these results are verified, preferably by multicenter studies that are structured to conform to accepted guidelines, such as those developed by the statistics subcommittee the National Cancer Institute and the European Organization for the Research and Treatment of Cancer [1, 2]. A biomarker that has been verified should then be validated

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4_25, © Springer Science+Business Media, LLC, part of Springer Nature 2019

in a large, prospective, multicenter study prior to its implementation. As this process requires significant resources, biomarker validation is usually supported by a commercial sponsor and conducted in conjunction with a clinical trial whose primary goal is to evaluate a novel therapy. However, in rare cases, biomarker validation can be conducted by international consortia using public funds, as is the case with validation of the Immunoscore, a prognostic immunohistochemistry-based biomarker for colon cancer based on localization of tumor-infiltrating T cells [3]. This validation study was published in 2018 and involved the analysis of tissue from 3539 patients in comparisson to the initial publication involving 415 patients published in 2006 [4]. The Immunoscore is a unique example of the significant scope, in both patient size and time, required to bring a biomarker from publication to use in personalized medicine. This chapter will outline this process using examples drawn from recent developments in cancer treatment using immunotherapy.

2 Evidence-Based Medicine

Clinical adoption of a new drug, practice guideline, or diagnostic assay relies on a very high level of evidence prior to implementation in the general population. The need to minimize harm, especially to overall healthy individuals, necessitates this high level of diligence. Many current changes to clinical practice are based on the established concepts of evidence-based medicine, where the evidence is placed in levels depending on its quality. Level 4 evidence is obtained from experts in the field and other individual experiences, whereas at the other end of the spectrum, level 1 evidence can only be obtained from meta-analysis of several randomized controlled trials. Clearly, level 1 evidence is impractical to obtain for biomarkers on their own; however, it is feasible to evaluate biomarkers as part of a larger clinical trial evaluating novel therapeutics. Counter intuitively, clinical trials that incorporate biomarkers are both more successful and more economical to conduct than traditional clinical trials [5]. In spite of this, the minority of clinical trials incorporate an integrated biomarker strategy. This combined approach of evaluating a therapeutic and diagnostic simultaneously is termed the Rx/Dx or theranostic approach.

An example of this approach is the development of a protein quantification assay for targets, such as programmed death-ligand 1 (PD-L1), which can be expressed on tumor cells in order to suppress the antitumor immune response. Thus, PD-L1 is an important target for therapeutic intervention with monoclonal antibodies, such as Pembrolizumab. Early evidence in the pivotal KEYNOTE-001 clinical trial where this drug was given to patients regardless of their PD-L1 expression level revealed that the

response rates were highest in patients in which at least half of the tumor cells expressed PD-L1, as determined by IHC [6]. This observation leads to subsequent studies designed to evaluate the assay performance in an effort to develop a validated companion diagnostic [7]. In spite of regulatory approval of this assay and others, there is a significant debate about both the analytical and clinical robustness of these assays. Recent studies that have compared four trial-validated IHC-based assays revealed discordance in 37% of patients between assays and scoring systems [8]. In spite of extensive research, consensus on the utility of PD-L1 IHC in prescribing anti-PD-L1 therapies remains controversial. The objective response rate (ORR) is 36% in patients with at least half of tumor cells expressing PD-L1, whereas the ORR falls to 19% in patients where between 1 and 24% of tumor expressed PD-L1 [9]. Therefore, the low positive predictive value (~33%) of PD-L1 as a predictive biomarker coupled with the high technical variability of the IHC-based assays used to quantify PD-L1 results in significant ambiguity for physicians in deciding whether a patient should receive anti-PD-L1 therapies. Specifically, even in patients with high PD-L1 expression, for every patient that responds, two will not respond, and most (70%) will experience treatment-related adverse events. To complicate issues further, even in patients with low (<25%) PD-L1 expression, one in five responds to anti-PD-L1 therapies. For patients who have failed to respond to other treatments, this off-label use may be a reasonable alternative. However, in the absence of predictive biomarkers, it is not possible to evaluate the risk-benefit ratio in these patients. Indeed, regulatory agencies recently updated prescribing information based on clinical trials and revealed that anti-PD-L1 treatment resulted in decreased survival in patients with very low (< 5% immune cell PD-L1-positive staining) when compared to chemotherapy [10]. Clearly, there is a need for additional predictive biomarkers that could be used to guide patient selection for anti-PD-L1 therapies.

Perhaps, it is unlikely that measurement of a single biomarker, even one as central as PD-L1, will enable accurate patient stratification in the context of a system as complex and dynamic as the immune system. If this is the case, multiplexed approaches, such as immunoproteomics, may have the potential to provide treatment selection assays that are both clinically and analytically robust. Morales-Betanzos et al. recently developed a quantitative proteomic assay that not only measures PD-L1 levels but also measures the extent of glycosylation [11]. They demonstrate that IHC staining intensity was lowest in samples containing high levels of glycosylation. This finding raises the concern that, in addition to the issue discussed above regarding variability in antibodies, instrumentation, and scoring systems used for IHC, target glycosylation may also affect the results obtained. From a treatment perspective, issues with staining of glycosylated antigen may partially explain

why a smaller but significant number of patients with apparently low target expression experience response compared to patient with high target expression. Fortunately, binding of the therapeutic antibodies to target PD-L1 is not likely affected by target glycosylation [12]. Nonetheless, it is likely that target glycosylation contributes to both disease progression and treatment efficacy.

The development of anti-PD-L1 and other checkpoint inhibitors is the subject of immense activity and has led to significant improvement in patient outcomes. These advances in the clinic underscore the complexity of the immune system and the potential utility of techniques such as immunoproteomics in aiding treatment optimization. In addition to PD-L1, the study by Morales-Betanzos et al. mentioned above included quantification of other immunoproteomic biomarkers, specifically PD-L2, IDO1, LAG3, CD40, TIM-3, and VISTA. These proteins all modulate T-cell antitumor activity. Although technical feasibility was demonstrated, the relevance or utility of these proteins as predictive biomarkers was not addressed, presumably since this is beyond the scope of this feasibility study of 22 patients. The discovery and verification validation of immunoproteomic biomarkers requires more extensive studies that expand both the range of biomarkers studied, in order to attempt to capture the complexity of the tumor–immune interface, and the analytical performance of the assay technologies themselves so that larger, sufficiently powered studies can be conducted. A recent study by Chevrier et al. examined 67 potential immunoproteomic biomarkers in tumors from 73 clear cell renal cell carcinoma (ccRCC) patients by mass cytometry. This technique shares many similarities with flow cytometry but has little to no overlap between the various channels. This increased specificity is afforded by the use of metals to tag antibodies instead of fluorophores. The study revealed the presence of 17 tumor-associated macrophage subsets and 22T-cell phenotypes, some of which are strongly associated with duration of progression-free survival [13]. The underlying technology, inductively coupled plasma mass spectrometry (ICP-MS), has been used for quantitative analysis of inorganic analytes for decades.

The two immunoproteomic studies mentioned above illustrate that multiplexed analytical tools can be used to reveal the complexity of the tumor-immune system interface and, although this complexity does impose significant demands on data analysis and interpretation, they can result in more accurate understanding, classification, and, eventually, prediction of outcomes in immunotherapy.

3 Statistical Challenges in the Validation of Immunoproteomic Biomarkers

A topic that is often overlooked in the design and interpretation of immunoproteomic biomarker studies is the unique statistical challenges that present themselves when we apply multiplexed

techniques such as immunoproteomics to the study of biological systems. Biomarkers, especially those associated with processes that are complex and low prevalence, are challenging to identify and very difficult to validate. For example, the search for diagnostic biomarkers is often performed using cohorts made up of equal numbers of cancer patients and controls. Even if these studies provide a sensitivity of 95%, which many do not, the actual positive predictive value drops to less than 20% if the disease prevalence is <1% [14], which fortunately is the case for most cancers. In other words, for every cancer detected, several false positives are encountered. Therefore, diagnostic biomarker studies conducted at prevalence rates that are different than that found in the population should not be extrapolated but rather verified in larger studies as discussed above.

Even for predictive biomarker studies, where prevalence is not an issue since all patients have a confirmed diagnosis, the high dimensionality of immunoproteomic measurements causes difficulty in identifying true biomarkers in the midst of a large background of false positives. Most researchers use *p*-values to indicate the strength of their biomarker panels. Many fail to correct these *p*-values for multiple hypotheses or ensure that the studies are sufficiently powered in the first place. Most studies cannot estimate power since the population variance is unknown. The practical result is that all too often, for studies involving small numbers of patients (e.g., <30), a *p*-value of 0.05 is likely obtained by chance and subsequent studies usually result in both a *p*-value and effect size that are very different from the original study [15].

4 Path Forward

In spite of the challenges in bringing an immunoproteomic biomarker from publication to practice, the need for novel biomarkers is growing as the number of approved therapies increases. Therefore, we conclude with strategies for increasing our ability to develop immunoproteomic biomarkers that are robust and can be implemented in the clinic.

4.1 Fitting with the Standard of Care

It is worthwhile to consider clinical implementation early on in the development of immunoproteomic biomarkers. While the analysis of tumor tissue and adjacent normal tissue does provide the most informative approach for comprehensive immunophenotyping in a research setting, the invasive nature of obtaining biopsies from solid tumors may make this approach challenging in many clinical settings. For example, for recurrent metastatic disease, a biopsy is recommended but is only obtained in the minority of patients in spite being associated with improved outcome [16]. Biopsies often do not yield tissue that is of sufficient quality or quantity suitable

for diagnostic assays. Furthermore, in the cases of immuno-oncology studies, serial sampling would likely be required due to the dynamic nature of the initial immune response and the subsequent durable, memory response that is likely required for curative therapies.

Liquid biopsies have the potential to fulfill the need for multiple, minimally invasive measurements that can be used to monitor the antitumor immune response. Frequent sampling of blood is part of the standard of care in many treatment regimens; therefore, feasibility of implementing blood-based biomarker strategies in immuno-oncology is high. This practicality is tempered by the fact that blood perfuses every organ and that there are strong homeostatic mechanism in place to regulate levels of many blood components. Therefore, there is a high potential for both false positives and false negatives due to, for example, significant discrepancies between blood-based biomarker levels and those found at the tumor site.

4.2 Technological Complexity

Validation of a biomarker does not guarantee that it will be used by clinicians, just as many approved drugs are commercial failures. Successful clinical adoption of novel technologies, whether they be therapeutic or diagnostic, typically requires significant improvement relative to the current standard of care. Even so, adoption will be hampered by tests or treatments that simply do not fit with current clinical practice. In the case of immunoproteomic techniques, familiarity with mass spectrometry is low, and this technology is not part of the clinical training of oncologists or the pathologists that make therapeutic and diagnostic decisions, respectively. Therefore, the path forward for adoption of immunoproteomic technologies must involve the development of methods that are accessible in terms of both technical complexity and result interpretation. Ideally, multivariate immunoproteomic methods that seek to compliment established univariate methods, such IHC and ELISA-based diagnostic assays, will be the most successful.

5 Conclusions

Immunoproteomic biomarkers are experiencing an explosive growth due to advancement in analytical technologies and the rapid growth of combination immunotherapies. These technologies, if applied to well-designed, large, collaborative studies, have the potential to result in tremendous gains for patients' prescribed immunotherapies. We hope that the steps and challenges in bringing immunoproteomic biomarkers from publication to implementation in personalized medicine that we have highlighted here are of use to those pursuing this challenge.

References

1. McShane LM, Altman DG, Sauerbrei W et al (2006) Reporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 100:229–235
2. Altman DG, McShane LM, Sauerbrei W, Taube SE (2012) Reporting recommendations for tumor MARKer prognostic studies (REMARK): explanation and elaboration. PLoS Med 9:e1001216
3. Pagès F, Mlecnik B, Marliot F et al (2018) International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 391:2128–2139
4. Galon J, Costes A, Sanchez-Cabo F et al (2006) Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 313:1960–1964
5. Falconi A, Lopes G, Parker JL (2014) Biomarkers and receptor targeted therapies reduce clinical trial risk in non–small-cell lung cancer. J Thorac Oncol 9:163–169
6. Rizvi NA, Garon EB, Leighl N et al (2015) Optimizing PD-L1 as a biomarker of response with pembrolizumab (pembro; MK-3475) as first-line therapy for PD-L1–positive metastatic non-small cell lung cancer (NSCLC): Updated data from KEYNOTE-001. J Clin Oncol 33(suppl)
7. Roach C, Zhang N, Corigliano E et al (2016) Development of a companion diagnostic PD-L1 immunohistochemistry assay for Pembrolizumab therapy in non-small-cell lung cancer. Appl Immunohistochem Mol Morphol 24:392–397
8. Hirsch FR, McElhinny A, Stanforth D et al (2017) PD-L1 immunohistochemistry assays for lung cancer: results from phase 1 of the blueprint PD-L1 IHC assay comparison project. J Thorac Oncol 12:208–222
9. Garon EB, Rizvi NA, Hui R et al (2015) Pembrolizumab for the treatment of non–small-cell lung cancer. N Engl J Med 372:2018–2028
10. USFDA. (2018) Center for Drug Evaluation, Research Drug Safety and Availability - FDA alerts health care professionals and oncology clinical investigators about an efficacy issue identified in clinical trials for some patients taking Keytruda (pembrolizumab) or Tecentriq (atezolizumab) as monotherapy to treat urothelial cancer with low expression of PD-L1. https://www.fda.gov/drugs/drug-safety-and-availability/fda-alerts-health-care-professionals-and-oncology-clinical-investigators-about-efficacy-issue. Accessed 05/22/2019
11. Morales-Betanzos CA et al (2017) Quantitative mass spectrometry analysis of PD-L1 protein expression, N-glycosylation and expression stoichiometry with PD-1 and PD-L2 in human melanoma. Mol Cell Proteomics 2017 16(10):1705–1717
12. Lee HT, Lee JY, Lim H et al (2017) Molecular mechanism of PD-1/PD-L1 blockade via anti-PD-L1 antibodies atezolizumab and durvalumab. Sci Rep 7:5532
13. Chevrier S, Levine JH, Zanotelli VRT et al (2017) An immune atlas of clear cell renal cell carcinoma. Cell 169 e18:736–749
14. Mausner JS, Kramer S (1985) Analytic studies. Mausner & Bahn epidemiology-an introductory text, 2nd edn. WB Saunders, Philadelphia, pp 154–194
15. Halsey LG, Curran-Everett D, Vowler SL, Drummond GB (2015) The fickle P value generates irreproducible results. Nat Methods 12:179–185
16. Shachar SS, Mashiach T, Fried G et al (2017) Biopsy of breast cancer metastases: patient characteristics and survival. BMC Cancer 17:7

INDEX

A

Adaptive immunity 3, 5, 10, 301
Adjuvants 20
Affinity purification 43, 168, 170, 175–176, 178, 254
Antibodies 2, 60, 80, 104, 119, 137, 153, 167, 183, 199, 216, 236, 248, 261, 288, 302, 310, 328, 333, 354, 372
Antigen discovery 32, 33, 181–197
Antigenicity 30, 31, 353–367
Antigen identification 27, 59–76, 120, 126, 134, 167–180, 195–196
Antigens 25, 59, 119, 138, 167, 182, 199, 213, 237, 246, 287, 327, 333, 354, 373
Autoantibodies 28, 30, 33, 60, 199–204

B

Bacteriophage lambda 181, 183, 186, 188
Bead array 79–92
Biomarker validation 372
Biomarkers 25, 28–30, 32, 34, 36, 39, 45, 60, 83, 96, 103–105, 120, 199, 259, 327–332, 371–376
Breakdown products 161

C

Cancer antigens 27, 34
Cancers ..... v, 11, 25, 60, 80, 96, 103, 199, 214, 235, 245, 259, 269, 327, 333
CD4+T cell epitope mapping 215, 217
Cell activation 16, 18, 215–217, 287, 314–316
Checkpoint 13, 21, 374
Chemical derivatization 247–249
Chemokines 6, 11, 35, 37, 79–92
Chimeric antigen receptor T cell therapies (CAR-T) 22
Crosslinking 167–180
Cytokines 2, 6, 8, 11, 13, 14, 16–18, 35–39, 45, 79–92, 287, 310, 315, 322, 340

D

Dissolution of immunoprecipitates 119
Druggable 310, 311, 317–321

E

Electroimmunoprecipitation 35, 119–120, 125, 126, 133
Epitopes 26, 133, 138, 161, 182, 200, 213, 245, 287, 302, 310, 328, 356
Evidence based medicine 372–374
Extravasation 310–312, 321

F

Flow cytometry 229, 374
Francisella 61, 288

G

GalNAc 200, 202–205, 209
Genomic DNA display 183
Glycopeptide analysis 279
Glycopeptide microarray 199–204
Glycosylation 30, 33, 38, 43, 154, 200–205, 208, 269, 281, 373

H

Hexapeptide 104, 106
High throughput screening 28–30
Human norovirus 137–151

I

IgSelect 155–159, 162, 163
Immobilized metal-affinity chromatography (IMAC) 44, 260–262, 264, 266
Immune complexes (ICs) 33, 34, 133, 333–338
Immune-informatics 45
Immune system ..... v, 1–23, 25, 30, 31, 40, 45, 235, 269, 270, 327, 328, 333, 341, 354, 373, 374
Immunity 2, 3, 8, 30, 36, 42, 137, 288, 301, 354
Immunoaffinity purification 41, 44, 235–242
Immunoassays 359, 364, 366
Immunoelectrophoresis (IE) 119–123, 126–128, 133, 134
Immunopeptidomics 43, 245, 247
Immunoprecipitates 119, 120, 125, 129, 133, 302, 305
Immunoprecipitation (IP) 119, 125, 133, 138, 245, 246, 248–251, 254, 255, 260, 261, 263, 266, 302, 304, 319
Immunoproteomics ..... v, 25–45, 59–76, 327–332, 353–367, 371–376
Immunoreactivity 60, 64, 67–69, 75
Immunotherapies ..... v, 21, 22, 32, 43, 44, 214, 260, 372, 374, 376

Kelly M. Fulton and Susan M. Twine (eds.), *Immunoproteomics: Methods and Protocols*, Methods in Molecular Biology, vol. 2024, https://doi.org/10.1007/978-1-4939-9597-4, © Springer Science+Business Media, LLC, part of Springer Nature 2019

Infectious diseases ..... v, 22, 29–31, 34, 42, 60, 103, 140, 184, 199, 328, 333
Inflammation..... 5, 8, 13–16, 36, 309–324
Influenza..... 22, 29, 216, 221, 227, 233, 353–367
Influenza A viruses 229–231
Innate immunity..... 2
Intact cells 340
Intact LC-MS 157, 159, 161
Integrative platform..... 327–332
Intercellular 11, 321

L

Liquid biopsies 376
Liquid chromatography tandem mass spectrometry (LC-MS/MS) 98, 154, 246, 250, 252–256, 260, 262, 265, 302, 304, 305, 314
Live vaccine strain (LVS)..... 62, 69, 288, 296
Low abundance protein enrichment..... 104, 105, 117
Low molecular weight proteome 95–101
Luciferase immunoprecipitation system (LIPS) assay..... 137–151
Luminex 79, 80, 82–84, 87, 89

M

Macrophages 2, 5, 6, 8, 10, 14, 15, 40, 44, 340–342, 347, 350, 374
Major histocompatibility complex (MHC) ..... v, 8, 26, 213, 235, 245, 260, 269, 287, 301, 328
Major histocompatibility complex (MHC)-associated glycopeptides 270, 273–277
Major histocompatibility complex (MHC)-associated peptide isolation..... 40, 41, 241, 250, 262, 271
Major histocompatibility complex (MHC)-binding peptides 214, 245
Major histocompatibility complex (MHC) epitope prediction..... 287–299
Major histocompatibility complex (MHC) ligandome..... 302
Major histocompatibility complex (MHC) ligands..... 288, 289, 301
Major histocompatibility complex (MHC) peptides..... 16, 26, 40–44, 235, 241, 245–256, 259–261, 277, 278, 282, 283, 296, 302, 304
Mass spectrometry (MS)..... v, 26, 59, 95, 104, 120, 154, 171, 202, 235, 245, 259, 270, 302, 313, 334, 339, 353, 374
Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)..... 202, 205, 339
Mild acid elution (MAE)..... 41, 245, 246, 249, 251, 255
Molecular modeling 269–283
Multiplex assay..... 87

N

Nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS)..... 311, 314, 316–317, 320, 323, 334–337
Neoantigens..... 43, 45, 214, 255, 256, 260, 328
NetMHC 302–305

P

Papain..... 334, 336, 337
Peptide elution..... 40, 41
Peptide library 108, 213–233
Phage-display library..... 181–197
Phosphorylation 196, 224, 259, 265
Post translational modification (PTM) 30, 33, 38, 43, 200, 201, 269
Protein depletion 108
Protein-protein interactions 39, 310, 311, 314, 315, 318–321
Proteomics..... v, 22, 25, 68, 95, 103, 154, 168, 256, 303, 309, 329, 333, 340, 373
ProteoMiner™ 103–117

R

Receptor identification 167
Renilla luciferase (Ruc)..... 138, 141–143, 146–149

S

Serological analysis of recombinant cDNA expression libraries (SEREX) 32, 33, 60, 199
Serological proteome analysis (SERPA)..... 27, 28, 59–76
Serum 13, 29, 60, 90, 96, 103, 132, 137, 154, 183, 206, 247, 260, 333, 340, 350
  biomarkers 39, 104
  stability 155
Sialylation..... 208
SpectMHC..... 302–306
Streptococcus pneumoniae..... 32, 183, 184, 186, 189–196
Surveillance 8, 36, 302, 328, 353–355, 366

T

Target prioritization 310
T-cell..... 8, 26, 213, 235, 260, 269, 287, 301, 311, 372
Therapeutics..... 28, 34, 42, 153–165, 214, 245, 302, 309–324, 372, 374, 376
Tularemia..... 61, 64, 288

V

Vaccines..... 1, 26, 60, 183, 199, 214, 245, 328, 353
Virus..... 29, 41, 137, 138, 216, 221, 229, 231, 353–359, 363, 366

W

Western blotting..... 25, 27, 59–61, 66–69, 74, 138, 156, 162–164, 246, 254, 263, 266, 320, 321

Y

Yeast display 216, 217, 226, 229–231