

Review

Assessment of sequence homology and cross-reactivity

Rob C. Aalberse*

*Department of Immunopathology, Sanquin Research at CLB, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands
Landsteiner Laboratory, Academic Medical Centre, 1066 CX Amsterdam, The Netherlands*

Received 14 July 2004; accepted 25 January 2005

Abstract

Three aspects of allergenicity assessment are discussed: IgE immunogenicity, IgE cross-reactivity and T cell cross-reactivity, all with emphasis on in-silico predictability: from amino acid sequence via 3D structure to allergenicity.

- (1) IgE immunogenicity depends to an overwhelming degree on factors other than the protein itself: the context and history of the protein by the time it reaches the immune system. Without specification of these two factors very few foreign proteins can be claimed to be absolutely non-allergenic. Any antigen may be allergenic, particularly if it avoids activation of TH2-suppressive mechanisms (CD8 cells, TH1 cells, other regulatory T cells and regulatory cytokines).
- (2) IgE cross-reactivity can be much more reliably assessed by a combination of in-silico homology searches and in vitro IgE antibody assays. The in-silico homology search is unlikely to miss potential cross-reactivity with sequenced allergens. So far, no biologically relevant cross-reactivity at the antibody level has been demonstrated between proteins without easily-demonstrable homology.
- (3) T cell cross-reactivity is much more difficult to predict compared to B cell cross-reactivity, and its effects are more diverse. Yet, pre-existing cross-reactive T cell activity is likely to influence the outcome not only of the immune response, but also of the effector phase of the allergic reaction.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Allergen; Allergenicity; IgE; Cross-reactivity

Contents

Introduction	S149
IgE immunogenicity	S150
IgE cross-reactivity.	S150
Cross-reactivity of T cells	S150
Conclusions	S151
References	S151

Introduction

The “allergenic potential” of a novel protein is a reflection of at least two distinct propensities:

- (1) to induce IgE antibodies (i.e., IgE immunogenicity, or “true” allergenicity). The interactions of the potential

* Fax: +31 20 5123170.

E-mail address: r.aalberse@sanquin.nl.

allergen involve among others the key cells of the immune system: antigen-presenting cells (APCs), various types of T cells and B cells;

- (2) to react with IgE antibodies induced by other proteins (i.e., cross-reactivity). The potential allergen can be active at two levels: influencing the immune response in a quantitative or qualitative way or, more common, interact with cell-bound IgE antibodies and trigger the effector phase of the allergic reaction.

These two are to a large degree independent. The prediction or evaluation of first feature, IgE immunogenicity, is the more difficult one by far. The second feature, cross-reactivity, is much more amenable to analysis, particularly if it is restricted to cross-reactivity at the antibody (B cell) level. A third feature, cross-reactivity at the T cell level is more complex, but will briefly be discussed as well.

IgE immunogenicity

The debate on “what makes an allergen an allergen” has stirred large clouds of dust, which has not settled yet. After a review of structural features of the allergens for which such information was available at the time (2000), I reached the conclusion that “allergens have no characteristic structural features other than that they need to be able to reach (and stimulate) immune cells and mast cells. Within this constraint, any antigen may be allergenic, particularly if it avoids activation of TH2-suppressive mechanisms (CD8 cells, TH1 cells)” (Aalberse, 2000). At the present time, I would have added “or other regulatory T cells, or the production of regulatory cytokines such as IL-10 or TGFβ”. Structural information on additional allergens has become available, but this does not really alter the message. A number of papers discuss algorithms to distinguish allergens from “non-allergens” based on searches in databases of protein primary sequences (Brusic and Petrovsky, 2003; Furmonaviciene and Shakib, 2001; Gendel, 1998a, 1998b, 2002; Ivanciuc et al., 2002, 2003a, 2003b; Li et al., 2004; Soeria-Atmadja et al., 2004; Stadler and Stadler, 2003). One major limitation of these papers is the ill-defined nature of the “non-allergens” and the position of the human homologues of allergens and non-allergens. Moreover, homology among allergens in the data base is not always taken into account.

Some of these in-silico analyses come up with potential would-be allergens. Obviously, it would be of major interest to find out how these predictions hold up in the real world, but, equally obvious, there is not much enthusiasm to test these predictions in humans.

IgE immunogenicity is, to a very large degree, determined by factors other than the protein itself. One very important factor is the way we, humans, are exposed to the protein (as a food, as an airborne protein, or as an injected protein). The discussion is largely initiated by concerns on the potential

allergenicity of novel proteins introduced into the food chain. Therefore, oral exposure is usually (even if implicitly) assumed which introduces the important aspects of thermal stability upon food preparation and digestibility in relation to gastric or intestinal fluids (Thomas et al., 2004). Yet, the databases used for in-silico allergenicity prediction usually include not only food allergens, but also allergens derived from airborne particles (pollen grains, mites, animal dandruff). Another factor is the environment in which the protein is introduced to the immune system, be it the food matrix or the intestinal milieu of the consumer. This may prevent or promote loss of structure and/or steer the immune system away from the TH2 pathway.

Prediction of IgE immunogenicity in a well-defined context (i.e., the same food matrix, the same immunization route, same host) thus depends largely on structural stability and only to a minor degree on intrinsic immunogenicity. The predictive value of even a combination of various assessment protocols (in-silico, in-vitro and animal models) is disappointingly low.

IgE cross-reactivity

IgE cross-reactivity can be much more reliably assessed by a combination of in-silico homology searches and in-vitro IgE antibody assays (Aalberse and Stapel, 2001; Aalberse et al., 2001). The in-silico homology search is unlikely to miss potential cross-reactivity with sequenced allergens. False warnings (i.e., protein pairs identified as homologous by structural analysis, but not obviously cross-reactive) will be found, but these can be subjected to serological cross-reactivity assays (which, obviously, requires access to a high-quality serum bank). So far, no biologically-relevant cross-reactivity at the antibody level has been demonstrated between proteins without easily-demonstrable homology.

By focusing on the relation between primary sequence and cross-reactivity, one important source of cross-reactivity is easily overlooked: post-translational modification. These processes are not hard-coded in the DNA sequence, but are determined by the host in which the protein is expressed.

Cross-reactivity of T cells

T cell cross-reactivity is much more difficult to predict than B cell cross-reactivity, and its effects are more diverse. So, it has conveniently been largely ignored. Yet, pre-existing cross-reactive T cell activity is likely to influence the outcome not only of the immune response, but also of the effector phase of the allergic reaction.

Most work on T cell cross-reactivity stems from research on autoimmune disorders. In this context, molecular mimicry is the term most commonly used. It specifically relates to T cell cross-reactivity between microbial antigens and auto-antigens. There is an impressive body of data

showing the striking lack of specificity (i.e., degeneracy) of many examples of the T cell–peptide–MHC interaction [see, for example, Ford and Evavold, 2004; Hunig and Bevan, 1981; Nino-Vasquez et al., 2004; Parnes, 2004; Scherer et al., 2004; Wilson et al., 2004; Wucherpfennig, 2004; Zhou and Hemmer, 2004]. Despite the well-established fact that longer peptides are involved in the T_H cell–MHC class II interaction compared to the T_C–MHC class I interaction, the specificity of the former (CD4⁺ T cell–MHC class II) interaction is even lower than that of the latter. Using larger combinatorial peptide libraries, it has been found that cloned CD4⁺ T cells can in some case be stimulated by peptides (in MHC class II context) without recognizable homology.

There are several messages in relation to IgE-mediated allergy, of which I will mention just two: (1) we have to be extremely careful in translating allergen-induced T cell activation (proliferation or cytokine production) into evidence of prior contact between the allergen and the immune system; (2) the accidental history (type of infections, etc) of the T cell branch of the immune system in an individual may markedly influence (positively or negatively) the outcome of a subsequent contact with, e.g., a food allergen that happens to generate a peptide that cross-reacts on the T cell level. This may sound as a re-phrasing of the hygiene hypothesis, but it is important to appreciate that in this case, the infectious agent acts in an immunologically specific (if degenerate) way, rather than by stimulating a TH2-suppressive milieu in a non-specific way.

Conclusions

- (1) Intrinsic allergenicity of a protein can be predicted only to a very limited degree and is strongly influenced by extrinsic factors.
- (2) Cross-reactivity at the level of the B cell or the IgE antibody is much more easy to predict (and the discussion on the value, or lack of value, of 6, 7 or 8 contiguous peptides is by now obsolete), with the restriction that the in-silico homology search can work only for allergens that are in the database. For allergens without known sequence (largely minor allergens), serum-screening protocols may be useful.
- (3) T cell cross-reactivity in relation to allergic sensitization has not (*apparently* “not” *should be included here*) been studied in sufficient detail (because it is a very complicated subject), but may have a major impact on IgE-dependent allergenicity.

References

Aalberse, R.C., 2000. Structural biology of allergens. *J. Allergy Clin. Immunol.* 106, 228–238.

- Aalberse, R.C., Stapel, S.O., 2001. Structure of food allergens in relation to allergenicity. *Pediatr. Allergy Immunol.* 12 (S14), 10–14.
- Aalberse, R.C., Akkerdaas, J., Van Ree, R., 2001. Cross-reactivity of IgE antibodies to allergens. *Allergy* 56, 478–490.
- Brusic, V., Petrovsky, N., 2003. Bioinformatics for characterisation of allergens, allergenicity and allergic crossreactivity. *Trends Immunol.* 24, 225–228.
- Ford, M.L., Evavold, B.D., 2004. Degenerate recognition of T cell epitopes: impact of T cell receptor reserve and stability of peptide: MHC complexes. *Mol. Immunol.* 40, 1019–1025.
- Furmonaviciene, R., Shakib, F., 2001. The molecular basis of allergenicity: comparative analysis of the three dimensional structures of diverse allergens reveals a common structural motif. *Mol. Pathol.* 54, 155–159.
- Gendel, S.M., 1998a. The use of amino acid sequence alignments to assess potential allergenicity of proteins used in genetically modified foods. *Adv. Food Nutr. Res.* 42, 45–62.
- Gendel, S.M., 1998b. Sequence databases for assessing the potential allergenicity of proteins used in transgenic foods. *Adv. Food Nutr. Res.* 42, 63–92.
- Gendel, S.M., 2002. Sequence analysis for assessing potential allergenicity. *Ann. N. Y. Acad. Sci.* 964, 87–98.
- Hunig, T., Bevan, M.J., 1981. Specificity of T-cell clones illustrates altered self hypothesis. *Nature* 294, 460–462.
- Ivanciu, O., Schein, C.H., Braun, W., 2002. Data mining of sequences and 3D structures of allergenic proteins. *Bioinformatics* 18, 1358–1364.
- Ivanciu, O., Schein, C.H., Braun, W., 2003a. SDAP: database and computational tools for allergenic proteins. *Nucleic Acids Res.* 31, 359–362.
- Ivanciu, O., Mathura, V., Midoro-Horiuti, T., Braun, W., Goldblum, R.M., Schein, C.H., 2003b. Detecting potential IgE-reactive sites on food proteins using a sequence and structure database, SDAP-Food. *J. Agric. Food Chem.* 51, 4830–4837.
- Li, K.B., Issac, P., Krishnan, A., 2004. Predicting allergenic proteins using wavelet transform. *Bioinformatics* 20, 2572–2578.
- Nino-Vasquez, J.J., Allicotti, G., Borrás, E., Wilson, D.B., Valmori, D., Simon, R., Martin, R., Pinilla, C., 2004. A powerful combination: the use of positional scanning libraries and biometrical analysis to identify cross-reactive T cell epitopes. *Mol. Immunol.* 40, 1063–1074.
- Parnes, O., 2004. From interception to incorporation: degeneracy and promiscuous recognition as precursors of a paradigm shift in immunology. *Mol. Immunol.* 40, 985–991.
- Scherer, A., Noest, A., de Boer, R.J., 2004. Activation-threshold tuning in an affinity model for the T-cell repertoire. *Proc. R. Soc. Lond., B Biol. Sci.* 271, 609–616.
- Soeria-Atmadja, D., Zorzet, A., Gustafsson, M.G., Hammerling, U., 2004. Statistical evaluation of local alignment features predicting allergenicity using supervised classification algorithms. *Int. Arch. Allergy Immunol.* 133, 101–112.
- Stadler, M.B., Stadler, B.M., 2003. Allergenicity prediction by protein sequence. *FASEB J.* 17, 1141–1143.
- Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., Van Ree, R., Woolhiser, M., Zawodny, J., 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* 39 (2), 87–98 (Apr.).
- Wilson, D.B., Wilson, D.H., Schroder, K., Pinilla, C., Blondelle, S., Houghten, R.A., Garcia, K.C., 2004. Specificity and degeneracy of T cells. *Mol. Immunol.* 40, 1047–1055.
- Wucherpfennig, K.W., 2004. T cell receptor crossreactivity as a general property of T cell recognition. *Mol. Immunol.* 40, 1009–1017.
- Zhou, D., Hemmer, B., 2004. Specificity and degeneracy: T cell recognition in CNS autoimmunity. *Mol. Immunol.* 40, 1057–1061.