Autoimmunity Reviews 11 (2012) 180–190

Contents lists available at ScienceDirect

Autoimmunity Reviews

journal homepage: www.elsevier.com/locate/autrev

Review

Computational analysis of high-density peptide microarray data with application from systemic sclerosis to multiple sclerosis

Michael Hecker [a,b,*], Peter Lorenz [a], Felix Steinbeck [a,c], Li Hong [a,d], Gabriela Riemekasten [e], Yixue Li [d], Uwe K. Zettl [b], Hans-Jürgen Thiesen [a,c]

[a] Department of Immunology, University of Rostock, Rostock, Germany
[b] Department of Neurology, University of Rostock, Rostock, Germany
[c] Gesellschaft für Individualisierte Medizin mbH (IndyMED), Rostock, Germany
[d] Chinese Academy of Sciences, Shanghai Institutes for Biological Sciences, Key Lab of Systems Biology, Shanghai, PR China
[e] Department of Rheumatology and Clinical Immunology, Charité Universitätsmedizin Berlin, Germany

ARTICLE INFO

Available online 18 May 2011

Keywords:
Peptide microarrays
Antibody epitope mapping
Bioinformatics
Multiple sclerosis
Systemic sclerosis

ABSTRACT

Auto-antibodies are implicated in the pathophysiology of various autoimmune diseases. High-density peptide microarrays incubated with human serum can detect antibody reactivities against thousands of peptides. This enables the identification of new auto-antigens and the determination of the parts of protein antigens (epitopes) that are recognized by antibody paratopes. We discuss the utility of peptide microarrays to investigate epitope-antibody-recognitions (EAR) from systemic sclerosis to multiple sclerosis. The technology can help to establish reliable diagnostic and prognostic biomarkers employing a combination of antigenic peptides. We describe the specifics of peptide microarray data and present bioinformatic methods for their analysis. Quality control, data pre-processing and the filtering of specific peptides are demonstrated on an example data set. Peptide microarrays representing 24 selected proteins by 3235 overlapping 15mer peptides were used to measure antibodies in serum of 10 patients with limited cutaneous systemic sclerosis (SSC) and 10 healthy blood donors. The data showed a sparse and skewed distribution, and we observed strong individual differences since many peptide sequences were bound by antibodies of only one serum sample. In the sera of the SSc patients, but not of the healthy controls, we found antibodies to three peptides MGPRRRSRKPEAPRR, TPTPGPSRRGPSLGA and GPSRRGPSLGASSHQ that share a similar sequence motif (GP-R/S-RR). These peptides map to two known linear epitopes at the N-terminus of centromere protein A (CENPA), demonstrating the utility of peptide microarrays. Presented experimental and bioinformatic approach can be applied in the same manner for multiple sclerosis research.

© 2011 Elsevier B.V. All rights reserved.

Contents

1. Introduction . . . 181
2. Experimental and bioinformatic methods . . . 182
 2.1. Study population . . . 182
 2.2. Preparation and staining of peptide microarrays . . . 182
 2.3. Data pre-processing . . . 183
  2.3.1. Image analysis . . . 183
  2.3.2. Background correction . . . 183
  2.3.3. Data truncation . . . 183
  2.3.4. Outlier correction . . . 183

Abbreviations: AA, amino acid; Ab, antibody; ACA, anti-centromere Ab; BSA, bovine serum albumin; CENPA, centromere protein A; CNS, central nervous system; CSF, cerebrospinal fluid; Ctr, healthy control; CV, coefficient of variation; dcSSc, diffuse cutaneous systemic sclerosis; ELISA, enzyme-linked immuno sorbent assay; HEp-2, human epithelioma type 2; ID, identifier; IgG, immunoglobulin G; lcSSc, limited cutaneous systemic sclerosis; MAID, MA plot-based signal intensity-dependent fold-change criterion; MS, multiple sclerosis; NMO, neuromyelitis optica; OCB, oligoclonal bands; PDB, Protein Data Bank; SSc, systemic sclerosis; SAM, significance analysis of microarrays; SD, standard deviation.

* Corresponding author at: Department of Immunology, University of Rostock, Schillingallee 68, 18057 Rostock, Germany. Tel.: +49 381 494 5891; fax: +49 381 494 5882.
E-mail address: michael.hecker@rocketmail.com (M. Hecker).

1568-9972/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.autrev.2011.05.010

2.3.5. Averaging . . . 183
2.3.6. Secondary Ab-specific signals . . . 183
2.3.7. Quality control . . . 183
2.3.8. Normalization . . . 183
2.4. Filtering specific antibody reactivities . . . 184
3. Peptide microarray results for systemic sclerosis . . . 184
4. Antibody profiling in multiple sclerosis . . . 185
5. Concluding remarks and perspectives . . . 187
Disclosure statement . . . 189
Take-home messages . . . 189
Acknowledgments. . . . 189
References . . . 189

1. Introduction

A pathological role of plasma cell-secreted auto-antibodies is widely accepted for many autoimmune diseases. Antibodies (Ab) are therefore of great importance for diagnostic purposes. High-density peptide microarrays are a fascinating way to measure Ab reactivities against thousands of peptides simultaneously (antibody profiling) [1–3]. Peptide microarrays can be applied, for instance, in the context of infectious and autoimmune diseases to identify novel autoantigens and determine specific Ab recognition sequences (epitopes). They also facilitate the systematic analysis of auto-Ab in oncology or humoral immune responses to vaccination procedures. Even if Ab reactivities are not always causally linked to a particular disease and may reflect just a bystander effect, they are still of outstanding importance for clinical diagnostics serving at least as surrogate markers.

In comparison to protein microarrays [4,5], the advantage of peptide microarrays is that even modified (e.g. phosphorylated and citrullinated) peptides are easily synthesized in a standardized, reproducible and cost-effective manner. In contrast, purified or recombinantly produced proteins are often difficult to obtain and require extensive quality control. Proteins are also more difficult to spot on microarray glass slides. Peptide arrays are more convenient to use, and can reveal the antigenic epitopes of a protein. However, as a drawback of this technique, many peptides are needed to adequately represent a protein, and most of these peptides will not be bound by Ab (resulting in data sparseness). Moreover, peptide arrays enable the detection of antibodies against linear epitopes, but not directly the identification of conformational epitopes formed by amino acids (AA) that are only in close proximity in the tertiary protein structure. Hence, both techniques, protein and peptide microarrays, rather complement each other.

Sophisticated bioinformatics is required to draw meaningful conclusions from the complex data sets generated by peptide microarray experiments. Computational methods are needed for detection and correction of systematic errors (which includes quality control), quantification of stochastic effects, data pre-processing and visualization, and selection of relevant peptides (filtering). Machine learning algorithms may then be applied e.g. to group or classify patients, which could have implications for medical care (digitized immune diagnostics). However, no standard workflow to evaluate peptide microarray data does currently exist, though different methods have been utilized (see Table 1 for an overview) [6–16]. In this work, we like to present basic bioinformatic methods that we found convenient to analyze such data.

We demonstrate our workflow on a data set where we used high-content peptide microarrays to compare IgG Ab reactivities in sera of 10 patients with limited cutaneous systemic sclerosis with those in 10 healthy control sera (Ctr). Systemic sclerosis (SSc) is a connective tissue disease characterized by microvascular damage and fibrosis of the skin and internal organs [17]. SSc is regarded as an autoimmune disease as auto-Ab against various nuclear antigens can be detected in the serum of patients. Anti-centromere Ab (ACA) and anti-TOP1 Ab are the classical Ab which segregate two patient subsets: ACA (e.g. anti-CENPA and anti-CENPB Ab) are associated with limited cutaneous involvement (lcSSc), whereas anti-TOP1 Ab are associated with diffuse skin involvement (dcSSc) [18,19]. In case of lcSSc, skin thickening is relatively restricted to the fingers and hands with less serious internal organ damage. Patients with dcSSc often rapidly develop severe lung,

Table 1
Literature overview on methods applied for peptide microarray analysis.

Study	Signal definition	Normalization	Filtering
Hueber et al. [6]	Median of the fluorescences of replicated spots	Scaling to a positive control (anti-lgM)	SAM corrected for multiple testing
Quintana et al. [7]	Not declared	Not declared	Mann–Whitney *U* test corrected for multiple testing
Gaseitsiwe et al. [8], Reilly et al. [9]	Log ratio of median spot and background fluorescence averaged over replicates	Intra- and inter-array by a linear mixed model learned on negative control spots	Signal threshold, signal must not be flagged or secondary Ab-specific, SAM
Nahtman et al. [10]	Log ratio of median spot and background fluorescence	Intra- and inter-array by a linear mixed model learned on all spots	Signal threshold, signal must not be flagged or secondary Ab-specific, SAM, fold change
Kidd et al. [11]	Median of the fluorescences of replicated spots, truncation of low signals	Not declared	SAM on log2 transformed data, corrected for multiple testing
Robinson et al. [12]	Median of the fluorescences of replicated spots, truncation of low signals	Scaling to a positive control (anti-lgG)	SAM on log2 transformed data, corrected for multiple testing
Quintana et al. [13]	Mean of log2 transformed background-subtracted fluorescences of replicated spots	Scaling to the median signal intensity of all antigens on the array	Mann–Whitney *U* test corrected for multiple testing
Robinson et al. [14]	Median of the background-subtracted fluorescences of replicated spots	Scaling to a positive control (anti-lgG)	Signal threshold (4-fold above mean signal intensity of negative control features)
Shreffler et al. [15]	Mean of the fluorescences of replicated spots	Scaling to a positive control, adjustment to the median signal intensity per chip	Log10 signal threshold (mean + 2 SD of all signals on the chip/on healthy control chips)
Neuman de Vegvar et al. [16]	Shifted median of the background-subtracted fluorescences of replicated spots	Scaling to a positive control (anti-lgG)	SAM on log2 transformed data, corrected for multiple testing, fold-change, signal threshold

The table lists ten antibody profiling studies that used microarrays containing about 200 peptides or proteins. Bioinformatic methods were categorized by signal specification, data normalization and filtering of relevant peptides. Ab = antibody, SAM = significance analysis of microarrays, SD = standard deviation.

renal or myocardial complications. In consequence, the prognosis and survival is worse in dcSSc than in lcSSc. However, despite their relation to disease course and progression, a definitive role for anti-nuclear Ab in the pathogenesis of SSc has not been shown so far. In this study, we screened the sera of the 10 lcSSc patients and 10 healthy blood donors for auto-Ab using microarrays containing 3235 different overlapping peptides derived from protein sequences of candidate autoantigens. As a result, bioinformatic analyses revealed the presence of serum Ab against the N-terminal tail of CENPA in all patients but in none of the controls.

Peptide microarrays can be employed to investigate many other diseases, including neurological autoimmune disease. Two prominent examples of very specific Ab biomarkers with diagnostic relevance in neurologic diseases are anti-AQP4 Ab for neuromyelitis optica [20] and anti-AChR Ab for myasthenia gravis [21]. We will review the proteins that are currently discussed as auto-Ab targets in multiple sclerosis (MS), an inflammatory disease of brain and spinal cord. We briefly outline the use of peptide microarrays to establish Ab signatures relevant for MS diagnosis and management, and suggest this technology as a powerful tool for clinical research.

2. Experimental and bioinformatic methods

2.1. Study population

Sera were obtained from 10 patients (all female; mean age: 60.6 ± SD 11.3 years) suffering from lcSSc (sample IDs lcSSc_01–lcSSc_10) treated in Berlin, Germany. Patients fulfilled the American College of Rheumatology (ACR) criteria for systemic sclerosis, assessed according to the European League Against Rheumatism (EULAR) Scleroderma Trials And Research (EUSTAR) criteria [17,22,23]. All patients were positive for ACA (tested by indirect immunofluorescence on HEp-2 cells) and anti-CENPB Ab (tested by line immunoassay), but not for anti-TOP1 Ab (tested by indirect immunofluorescence on HEp-2 cells). The study was approved by the local ethics committee (Charité-Universitätsmedizin Berlin, Germany). Patients signed an individual informed consent. As a control group, we analyzed sera of 10 healthy blood donors (sample IDs Ctr_01–Ctr_10; 6 females, 4 males; mean age: 49.4 ± SD 13.3 years) kindly supplied by Dr. Renate Claus (University of Rostock, Germany).

2.2. Preparation and staining of peptide microarrays

High-density microarrays were custom-made at JPT Peptide Technologies GmbH (Berlin, Germany). The peptides spotted on the microarrays (standard microscope slides, 76 × 26 mm) were selected on the basis of 24 autoimmune disease-associated (potentially antigenic) proteins (Table 2). All peptides derived for a certain protein were 15mers covering the whole protein sequence with 11 AA overlap. Details on the peptide synthesis and spotting can be found elsewhere [24]. On each microarray, the whole set of peptides was spotted in triplicate as three subarrays. In total, each peptide microarray contains about 10,000 spots with a diameter of 150 μm. Some spots carried Cy3-fluorescent dye to ease spot alignment after scanning (guide dots). Arrays were blocked immediately after spotting with bovine serum albumin (BSA) by the manufacturer JPT.

Microarray staining was performed at room temperature in a chamber formed by a 2-slide sandwich made by first adding 3 mm wide and 1 mm thick plastic spacers to either end of the microarray slide and then putting a new clean slide on top of the spacers. The staining protocol was previously described in detail by Lorenz et al. [24] and is summarized in the following. The first step was incubation with 500 μl of the primary Ab preparation, i.e. patient serum diluted 1:100 in TTB buffer (50 mM Tris/HCl pH 7.4 and 150 mM NaCl, 0.05% Tween 20, 0.1% BSA). This dilution of the sera was found appropriate to detect also low affinity and low abundance IgG-type Ab without raising the background signals. After 4 h at room temperature the slide sandwich was disassembled and the microarray consecutively washed three times with TTB and once with distilled water. Residual fluid was removed by a centrifugation step (1 min at 1000 g). Secondary incubation was done like the primary Ab incubation in the slide sandwich, however, incubation time was only 2 h and performed in the dark. Secondary Ab consisted of 500 μl of Zenon goat anti-rabbit IgG (0.2 mg/ml, Fab fragments) conjugated to Alexa Fluor 647 (Invitrogen Z25308). Thereafter, microarray slides were washed as described above with additional two washes with distilled water. Fluorescence signals were acquired immediately afterwards using a

Table 2
Proteins represented by peptides on the microarrays.

Official symbol	Swiss-Prot ID	Entrez gene ID	Official full name	Protein length	Peptides
APOH	P02749	350	Apolipoprotein H (beta-2-glycoprotein I)	345	84
CENPA	P49450	1058	Centromere protein A	140	33
CENPB	P07199	1059	Centromere protein B, 80 kDa	599	147
CENPE	Q02224	1062	Centromere protein E, 312 kDa	2701	624
EXOS10	Q01780	5394	Exosome component 10	885	219
FBL	P22087	2091	Fibrillarin	321	78
H1F0	P07305	3005	H1 histone family, member 0	149	46
H1FX	Q92522	8971	H1 histone family, member X	213	51
HIST1H1A	Q02539	3024	Histone cluster 1, H1a	215	66
HIST1H1B	P16401	3009	Histone cluster 1, H1b	226	76
HIST1H1C	P16403	3006	Histone cluster 1, H1c	312	81
HIST1H1D	P16402	3007	Histone cluster 1, H1d	221	83
HIST1H1E	P10412	3008	Histone cluster 1, H1e	219	87
HIST1H4E	P62805	8367	Histone cluster 1, H4e	103	41
HIST1H4G	Q99525	8369	Histone cluster 1, H4g	98	22
KIF11	P52732	3832	Kinesin family member 11	1056	258
MPO	P05164	4353	Myeloperoxidase	745	184
POLR3A	O14802	11128	Polymerase (RNA) III (DNA directed) polypeptide A	1390	331
POLR3B	Q9NW08	55703	Polymerase (RNA) III (DNA directed) polypeptide B	1133	281
PRTN3	P24158	5657	Proteinase 3	256	62
SSB	P05455	6741	Sjogren syndrome antigen B (autoantigen La)	408	100
TOP1	P11387	7150	Topoisomerase (DNA) I	765	189
TRIM21	P19474	6737	Tripartite motif-containing 21	475	116
TROVE2	P10155	6738	TROVE domain family, member 2	538	132

A total of 24 proteins was included in this analysis. “Protein Length” gives the number of amino acids for each protein sequence. Column “Peptides” provides the number of peptides corresponding to the respective proteins.

Molecular Devices Axon Instruments GenePix 4000B scanner with 532 and 635 nm laser excitation sources. Two scans at 10 μm resolution, one with 523 nm to read the Cy3 guide dots, and one with 635 nm to read the signals due to staining (Cy5), were performed at 16 bit depth.

According to this protocol, 20 peptide microarrays (10 lcSSc and 10 Ctr samples) were stained and scanned in four independent experimental runs profiling the Ab reactivities of I) lcSSc_02–03, II) lcSSc_01, Ctr_04, Ctr_07–09, III) lcSSc_04–05, lcSSc_10, Ctr_01–03, Ctr_06, and IV) lcSSc_06–09, Ctr_05, Ctr_10. In each run, one additional microarray was prepared as a negative control (TTB buffer plus secondary Ab, Buffer_1–4).

2.3. Data pre-processing

2.3.1. Image analysis

Each microarray image was processed with the GenePix Pro software (Molecular Devices) to calculate the median Cy5-fluorescence intensity for each peptide spot and its respective local background (i.e. the area surrounding a spot). The software also assigned flags to mark spots with signals lower than the background or spots that have an abnormal shape. The results of this image analysis were exported as GenePix Results (GPR) files.

2.3.2. Background correction

For each peptide spot we subtracted the background from the spot intensity to make the data more comparable between the (sub)arrays. This calculation was justified by the fact that higher non-specific spot intensities strongly correlate with higher background intensities (that occur e.g. in smears on a microarray). The ratio of spot intensity and background as used by others [8–10] was considered less adequate as background intensities are usually low and noisy. Therefore, dividing spot and background intensity would amplify stochastic effects and may rather distort the data.

2.3.3. Data truncation

In general, the signal intensity of most of the spots is quite low (data sparseness), often even lower than the background. The latter observation can be explained by the different surface of the glass slide in areas where peptides are spotted compared to the spot-surrounding areas that are just blocked by BSA. However, as it is not meaningful to operate with negative signals, we set the signal of a peptide to 1 if the fluorescence was higher in the background than in the spot (truncation).

2.3.4. Outlier correction

When comparing the signals of the three subarrays for an individual microarray one can usually find some outliers. An outlier is defined if the signal of one spot of the triplicate is notedly higher than the other two spot signals. Such outliers can be explained by local specks on the microarray and could considerably affect some data points. To detect such outliers we computed the coefficient of variation (CV) for each of the triplicates in the data set. Triplicates that have a CV higher than the 0.95-quantile of all CV were defined as outliers. Each outlier spot was then flagged and corrected by setting the spot intensity to the mean intensity of the other two spots of the triplicate.

2.3.5. Averaging

Finally, to end up with one signal value for each peptide and array, we used the mean of the triplicate spot intensities after background subtraction, truncation and outlier correction. As a rough rule of thumb, a signal below 500 can be regarded as uncertain and a signal below 200 indicates a spot that is not bound by Ab.

2.3.6. Secondary Ab-specific signals

Remarkably, the buffer control microarrays showed some high signal intensities as well. This is supposed to be an effect of the secondary Ab used for the staining that presumably binds to some peptides directly. This implies that in order to extract specific peptide signals from the data, we need to filter peptides a) by signal intensity and b) by demanding that the signal is significantly higher in the microarrays prepared with sera than in those prepared with buffer only instead (see Section 2.4).

2.3.7. Quality control

As mentioned before, some arrays are characterized by incidental smears that influence the data. Therefore, a quality control step should be considered to exclude microarrays of low quality from further analyses. We conceived a simple and effective measure to evaluate the quality of a microarray. For each array, we computed the correlation coefficients according to Pearson and Spearman for each pairwise combination of the 3 subarrays on the background-subtracted and truncated spot intensities of those spots that are not flagged in both subarrays to be compared. The quality control measure (consistency) is then defined by the average of these six correlation coefficients. A consistency lower than 0.7 indicates microarrays of poor quality.

2.3.8. Normalization

Normalization methods aim to adjust the data for systematic errors which arise from experimental and not biological variation. This could make the data of different peptide microarrays more comparable. However, according to our experience, the reproducibility of and similarity between microarrays is in general acceptable, in particular for arrays stained within one experimental run. Eventually, the inherent stochastic noise in these data might outweigh systematic errors. Hence, we did not further transform the data. We would rather emphasize that some normalization methods may be even more prejudicial than beneficial. Most of them have been initially developed for gene expression microarray data based on assumptions that are not applicable for the analysis of peptide microarray data. Differences between peptide and DNA microarray data are (as later demonstrated): a) the signal distribution is much more skewed and sparser for peptide arrays, i.e. for the vast majority of peptides the signal is low and only few peptides show a strong signal; b) in peptide array data sera of two patients can recognize very different peptides, and there might be no peptide that presents a high signal in all samples of a study. In comparison, in DNA microarray data many genes are usually similarly expressed across different samples (e.g. of a certain tissue or cell culture); c) for gene expression data the amount of overall signal is often assumed to be approximately similar in each array. For peptide arrays, this assumption is not justified as one sera might show reactivities against hundreds of peptides and another one to only few of them; and d) stochastic deviations are significantly higher in peptide array data even if they can be reduced by background subtraction, outlier detection and averaging over the triplicates.

Due to these data characteristics, it is difficult to reliably estimate the parameters of a normalization algorithm. Global scaling and other more sophisticated transformations thus seem inappropriate for peptide microarrays. Sources of systematic variation between independent experimental runs might be a) technical deviations in the concentration of Fab fragments employed for secondary Ab incubation and b) use of different batches of these polyclonal and polyspecific secondary Ab. Hence, it is important to prepare a buffer control microarray for each experiment. We further hypothesize that sample differences in the serum-constituents (other than IgG) contribute to some variation in the data as well. However, these effects cannot be corrected by normalization without risking the introduction of artifacts. It should be further noted that anti-IgG spots that were utilized for normalization by others (see Table 1) were not applied for this purpose in this study because of their often irregular spot geometry, saturation effects in these spots and hence lack of reproducibility.

2.4. Filtering specific antibody reactivities

As the obtained peptide signals have a skewed distribution and are subject to fluctuations, we implemented a semi-quantitative analysis to filter a subset of specific peptides. We distinguished two filtering steps: Firstly, we selected the peptides for which the signal of at least one serum-stained microarray is strong and significantly higher in comparison to the corresponding buffer control conducted in the same experimental run. Secondly, the remaining peptides were filtered according to their capacity to separate groups of patients.

The first filtering takes into account the amount of signal as well as the reactivities of the buffer control microarrays (stained only with buffer and secondary Ab). Often relatively few peptides survive this filtering, i.e. only some peptide signals usually turn out to be serum-specific for at least one individual sample. The method to identify peptides with serum-specific signals was derived from a method originally developed by our group to select differentially expressed genes in DNA microarray data [25]. This so-called MAID filtering method is briefly explained in the following.

To detect a peptide as specifically bound by serum Ab one can e.g. use two filtering criteria: a) the signal intensity for this peptide (e.g. signal >500) and b) the fold-change (that is the ratio) of this signal and the respective peptide signal obtained for the buffer control (e.g. fold-change >5). However, fixed thresholds for both do not adequately reflect the characteristics of the data. MAID instead implements a smooth combination of both criteria by defining a signal intensity-dependent fold-change (MAID score). Particularly, the method considers that the variability in the fold-changes increases as the measured signal intensity decreases because relative errors are higher at lower intensities [26]. The filtering algorithm is based on the MA plot visualization. For each of the 20 serum-stained peptide arrays an MA plot was produced that compares each array versus the buffer control array of the same experimental run. The algorithm then calculated separately for each MA plot a set of data points (S) representing the intensity-dependent variability. For S we used the 90%-quantiles of the M values which were computed in a moving window (in direction of A) comprising 100 peptide signals. Afterwards, the calculated data points in S were smoothed and an exponential function $f(x) = a^* e^{(-b^* x)+c}$ was fitted to them by a non-linear least squares estimation, constraining a, b and c to be positive. With the exponential function for each particular MA plot, we computed a MAID score for each peptide signal by dividing the M values by f(A). Peptides were finally defined as serum-specific if no more than one of the triplicate spots was flagged by GenePix Pro and if the (background-subtracted, truncated, averaged) signal resulted in a MAID score >1.0. To our experience, a lower cut-off can be acceptable but we have here chosen this quite stringent filtering.

On the data set reduced by this first filtering we applied a second filtering to select group-specific peptides. We compared the two cohorts, lcSSc patients and healthy blood donors, seeking for peptides that have a serum-specific signal for several patients but not for the control samples. To statistically evaluate the relevance of a peptide to distinguish the two groups in this rather qualitative (categorical) manner, we built a 2×2 contingency table for each peptide with the numbers of individuals who are lcSSc and Ab−, lcSSc and Ab+, Ctr and Ab−, and Ctr and Ab+, where Ab+ denotes that the peptide was filtered as serum-specific. Then we computed a p-value for each contingency table using Fisher's exact test that is testing the null hypothesis that the proportions in the groups are the same. Peptides with a raw p-value<0.01 were defined as group-specific. However, it should be noted, that this cut-off is also quite stringent. In some situations we might miss some relevant peptides. A higher p-value cut-off should be acceptable, e.g. if only a relatively small patient subpopulation shows Ab reactivities against a certain peptide, or if patients have different Ab recognizing different epitopes of the same protein. We further like to point out that when applying statistical tests (e.g. SAM [27]) directly on the signals (i.e. without categorization for serum-specificity) difficulties may arise if the variances in the data are extremely high (see Section 3). Log-transforming the data beforehand could to some extent overcome this issue. However, as these tests consider the distribution but not the amount of signal, one should be aware that peptides with very low signals in all samples might survive the filtering procedure by chance.

All analyses were performed in the R environment for statistical computing accessed through a local webserver maintained at the Institute of Immunology (University of Rostock, Germany).

3. Peptide microarray results for systemic sclerosis

Fluorescence scans of all 24 peptide microarrays (10 lcSSc, 10 Ctr and 4 buffer controls) were processed as described above. On each microarray, 24 proteins (Table 2) were represented by 3235 15mer peptides spotted in triplicates. Fig. 1 shows an exemplary microarray image and detailed views. The data are sparse and skewed: Most of the spots (65.11%) were flagged by the GenePix Pro software. Background-subtraction was applied to adjust for variation in the background. Background was higher than the spot fluorescence for 62.20% of the spots. Negative spot signals were truncated (i.e. set to 1) and outliers were removed. For instance, an outlier was found on the microarray stained with serum from Ctr_10: a peptide from CENPB (261–275) showed a signal of 19,891 in one subarray but a signal of 1 in the other two subarrays (CV = 1.73), and the data were corrected to be 1 in all three subarrays. For the serum-stained microarrays, only 9.99% of the peptide signals (after averaging the triplicates) were ≥200 and thus significantly higher than the background. However, all four buffer control arrays showed some peptides with a signal >1000.

In one experimental run, the buffer control microarray (Buffer_4) was excluded during quality control due to its low consistency of 0.3. Therefore, it was replaced by Buffer_1 for the identification of serum-specific peptides with the MA plot-based filtering algorithm. Fig. 2 shows an exemplary MA plot and visualizes the MAID filtering method. The observation that several peptides with A value >10 have an M value <0 at least for some samples might indicate a serum-effect, i.e. that the composition of individual serum impacts on the secondary Ab reactivities, which was already speculated by others [10]. Relatively few peptides passed the filtering: 488 out of the 3235 peptides were filtered as serum-specific for at least one of the 20 microarrays. Between 12 and 103 peptides were filtered per sample demonstrating the individual diversity in the number of peptides bound by Ab. Although the MAID score cut-off was chosen quite high, we found this a typical result for peptide microarray data. Many peptides (322 out of 488) showed a serum-specific signal only in one individual sample. Only few peptides had a high signal for several individuals, and not even one peptide was filtered to be bound by Ab for 15 or more of the 20 sera. We visualized the intensities of the filtered peptides as a heatmap in Fig. 3. It shows the sparse distribution of the data and the strong individual differences.

Three peptides were finally filtered as lcSSc-specific by Fisher's exact test. MGPRRRSRKPEAPRR from CENPA (1–15) was serum-specific in 9/10 patients and none of the healthy controls (p-value = 0.00006) and thus ranked the highest. Fig. 4 shows the data for this peptide and the huge diversity in the magnitude of the signals across different samples. Two other overlapping peptides from CENPA were filtered: GPSRRGPSLGASSHQ (25–39) and TPTPGPSRRGPSLGA (21–35). For those, none of the Ctr samples had high signals, but 7/10 (p-value = 0.002) and 6/10 (p-value = 0.005) of the patients, respectively. The three lcSSc-specific peptides all exhibit a similar sequence motif (GP-R/S-RR). They match to two known antigenic regions (epitopes) at the N-terminal part of the CENPA protein, which have been already identified by immunoblot and ELISA [28,29]. A recent publication proposed that the CENPA epitope 2 is homologous to the transcription factor FOXE3, which may hence either be targeted by the

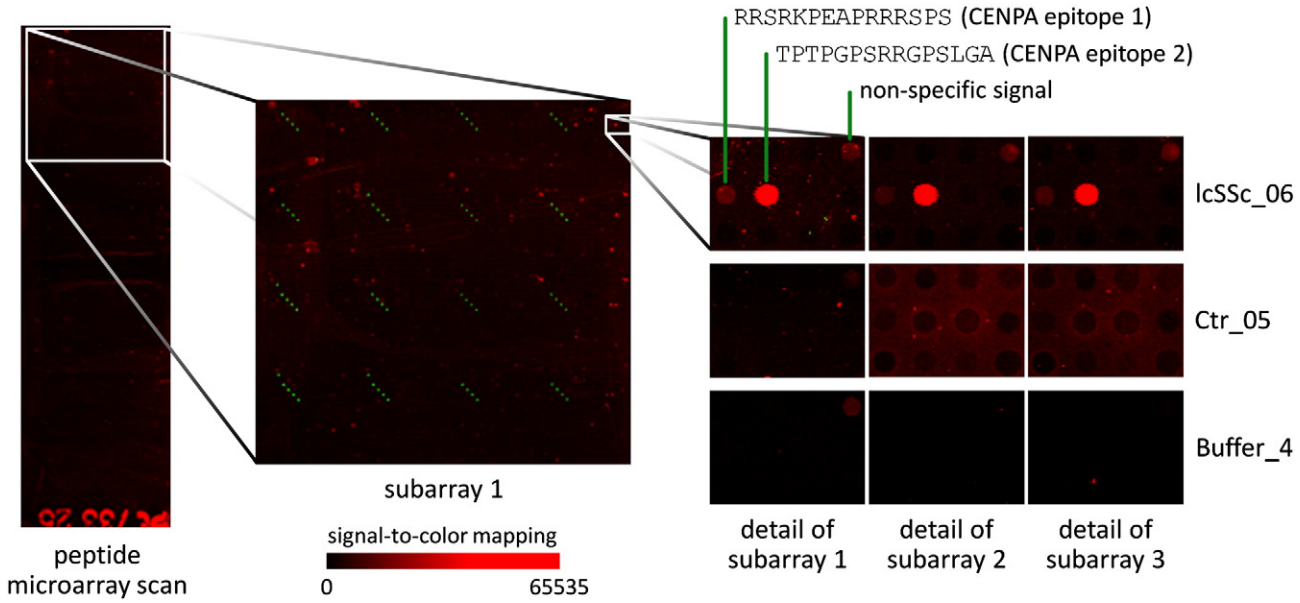

Fig. 1. Fluorescence image of a peptide microarray and detail views. Each glass slide consisted of three replicate subarrays (top, middle, bottom) and in each subarray antibody reactivities against 3235 peptides derived from 24 proteins (Table 2) were measured. The Cy5 fluorescence in the red channel was evaluated to assess whether a spotted peptide is bound by Ab. Cy3 signals (green channel) of the guide dots can also be seen in the subarray image. Detailed views on the right show a region of 1 mm width for all three subarrays of a microarray stained with serum of a lcSSc patient and of a healthy control, respectively, as well as of a buffer control array. Overall, the signal reproducibility among the subarrays and the shapes of the peptide spots were satisfactory, but we noted elevated backgrounds in some regions (subarrays 2 and 3 for Ctr_05). Two 15mer peptides from CENPA showed high signals for lcSSc_06: RRSRKPEAPRRRSPS (5–19) and TPTPGPSRRGPSLGA (21–35). In the upper-right of the region there is a spot with some fluorescence on both serum-stained microarrays as well as on the array only stained with buffer and the secondary Ab.

Ab or involved in priming them [30]. Table 3 provides the data for the three filtered peptides. As shown, all patients had a serum-specific signal for at least one of the peptides. However, three patients (lcSSc_01, lcSSc_04, and lcSSc_08) were only Ab-positive to CENPA epitope 1, while one patient (lcSSc_10) was only positive to CENPA epitope 2. Fig. 5 is a compact visualization (which we refer to as VisualMap) of the CENPA peptide signals determined for all 24 microarrays. It discloses the position where Ab potentially bind the protein (linear epitope mapping) as well as the so far only partly (AA 59–134) elucidated CENPA 3D structure [31]. Two regions (69–91 and 97–123) show Ab reactivities for both serum groups, lcSSc and Ctr. Both regions overlap with loops in the tertiary structure of the protein (L1, 80–87 and L2, 116–123). Possibly, reactivities of the secondary Ab might have caused elevated signals, since the buffer control microarrays showed signals above 300 for these two regions. Yet, this example illustrates how informative it is to know the true epitope sequence of a protein (e.g. in diagnostics), otherwise results might be misleading.

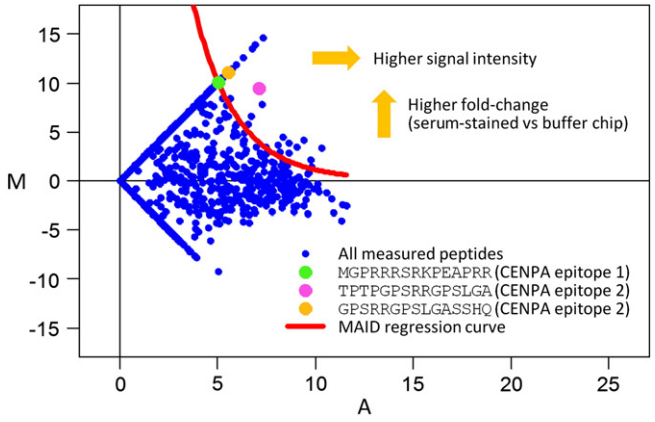

Fig. 2. Exemplary MA plot visualizing the filtering method. This MA plot compares the peptide signals obtained for lcSSc_03 and the corresponding buffer control (Buffer_1). For each peptide, A (x-axis) is the mean log2 signal of both microarrays and M (y-axis) is the log2 fold-change. A positive M value thus means that the peptide signal is higher in the serum-stained than in the buffer control array. Since the signals were truncated to be ≥1.0, the A value is ≥0 for all peptides. The MA plot reveals significant stochastic variation for low signal intensities. To identify peptides specifically bound by serum Ab of lcSSc_03, an exponential function (MAID regression curve, red line) was fitted to the signal intensity-dependent variability in the data. In this case, 21 peptides survived this filtering (blue dots above the red line) including the three peptides of CENPA epitope 1 and 2 (Table 3).

For comparison, we analyzed the 3235 peptides using Mann–Whitney *U* test and significance analysis of microarrays (SAM), respectively. The Mann–Whitney *U* test (raw *p*-value<0.001) identified four peptides with differential signals between both groups, including the three CENPA peptides discussed above. However, the fourth peptide showed a signal of <30 in all microarrays and is therefore considered to be a random result. SAM (raw *p*-value from permutations <0.001) filtered nine peptides, none of them had a signal >200 in any peptide array hence the data are noisy and not reliable. Because of the high variation in the signal intensities for GPSRRGPSLGASSHQ and TPTPGPSRRGPSLGA in particular (Table 3), SAM failed to identify these peptides (*p*-value>0.1). However, when we applied SAM on the log2 transformed peptide signals, it delivered the three peptides from CENPA epitope 1 and 2, as well as five other peptides of which three had a maximum signal of only around 200.

4. Antibody profiling in multiple sclerosis

Peptide microarrays can be subjected to the investigation of many other autoimmune diseases, including neurological diseases. Multiple sclerosis (MS) is the most common progressive and disabling neurologic disease of young adults. In people with MS, lesions accumulate in the brain and spinal cord and damage the myelin covering of nerve fibers. Inflammation and the loss of myelin cause disruption to nerve transmission and thus affect many functions of the body. The inflammatory processes are mainly driven by T and B cells that enter the central nervous system (CNS) through the damaged blood–brain barrier [32–34].

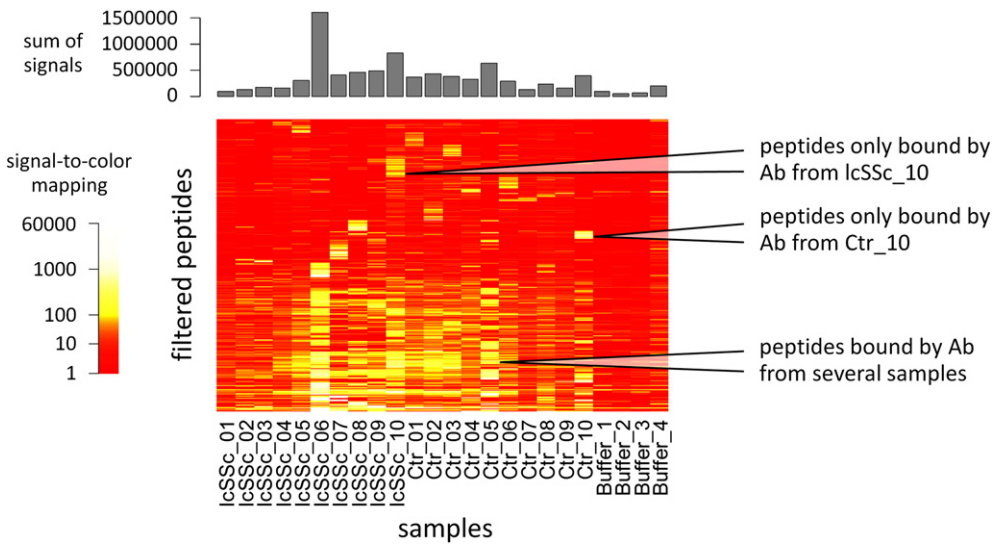

Fig. 3. Graphical representation of the reduced data set. Heatmap of signals from 488 peptides that were filtered by the criterion to be bound by serum Ab for at least one of the 20 samples. Very distinct individual Ab reactivities can be seen for patients and healthy controls likewise. Many peptides show a serum-specific signal only for one sample (upper half). From this visualization, there is no apparent difference between lcSSc and Ctr. Signals are generally low in the buffer control microarrays for the presented peptides. The sum of peptide signals was calculated for each column and is shown on top.

A role for Ab in MS is supported by observations of intrathecal Ab synthesis, oligoclonal B cell expansion in the cerebrospinal fluid (CSF), the frequent detection of immunoglobulin deposition in MS lesions and by the fact that specific Ab can induce demyelination and axonal loss in animals [35]. Oligoclonal bands (OCB) in the CSF are detectable in about 85% of MS patients and are therefore an important diagnostic tool [36]. OCB also suggest that only a limited number of B cell clones contributes to the elevated Ab levels. At present, there is a renewed interest in autoreactive B cells and auto-Ab in MS due to the positive effects with B cell depletion therapies on disease activity [37,38]. However, it is poorly understood how B cells are stimulated in the pathogenesis of MS. This activity could be induced by antigens (e.g. from either self or foreign proteins) or possibly represents just a random bystander effect during ongoing inflammation processes. According to the latter, a general activation of normal immune system

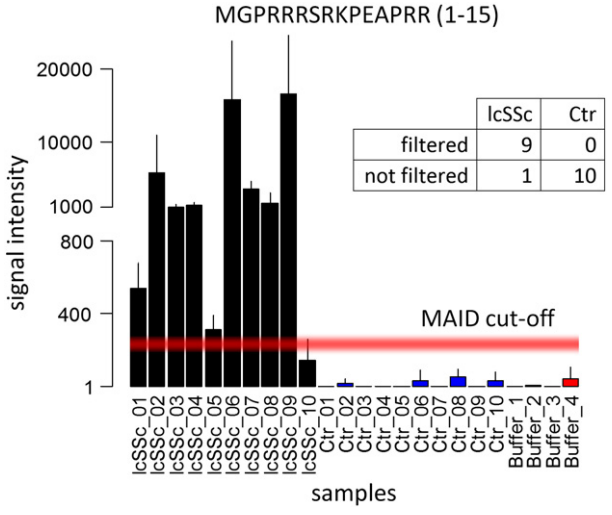

Fig. 4. Bar chart for the top lcSSc-specific peptide. Group-specific filtering indicated that the peptide MGPRRRSRKPEAPRR (1–15) was specifically recognized by Ab of nine patient sera (Fisher's test p-value = 0.00006). The data are skewed and peptide signals differed a lot between the patients: from less than 500 to about 15,000. The signal for the serum sample of lcSSc_10 did not exceed the MAID score cut-off. Nevertheless, it was still higher than the respective signals on all control-sera and buffer microarrays. Lines above the bars depict the standard deviations computed for each triplicate.

processes may attract Ab producing cells and trigger stronger Ab responses as an epiphenomenon. Undesired humoral immune responses are also observed in the treatment of MS. Some patients develop neutralizing Ab against the drugs used for therapy (e.g. interferon-beta [39] and natalizumab [40]). These Ab reactivities are believed to lessen therapeutic benefits and therefore should be assessed in regular intervals during therapy.

Many Ab (mainly IgG but also other isotypes) have already been proposed for MS diagnosis, prognosis and monitoring [34,35,41,42]. Table 4 reports the auto-Ab protein targets that are under discussion [43–66]. Further Ab specificities have been associated to MS, e.g. against lipids [67], glycans [68] and viral proteins [69] (important in the context of molecular mimicry). The production of myelin-specific Ab is often seen as one of the most important ways how plasma cells may contribute to the pathogenesis of MS. However, though several potential Ab markers have been identified, so far not one single known Ab specificity has been detected in the serum or the CSF which can definitely distinguish MS patients from healthy controls or other diseases. In fact, in the studies listed in table 4 only a subset of patients shows elevated Ab levels against certain proteins. Some of the Ab measures might not be specific enough. Moreover, in some instances, findings could not be fully reproduced by other researchers. The latter might be explained by the use of different assays measuring Ab reactivities against different parts or conformational structures of a protein. In addition, other factors can play a role as well: a) the microenvironment analyzed (blood, brain, spinal cord), b) the age of the patients [70] or c) the use of different statistics to analyze data with mostly skewed distribution.

Possibly, many MS variants with different (immuno)pathologies exist. This assumption could explain why single markers have only moderate sensitivity and specificity. To overcome this issue, different Ab markers from sera (systemic) and CSF (local) need to be combined into a biomarker panel. Peptide microarrays can be an excellent tool to assess such a broad panel of different auto-Ab for MS. The application of microarray antibody profiling can therefore lead to more specific diagnostic and prognostic markers, and may provide new hypotheses regarding underlying disease processes.

Peptide microarrays have already been applied to investigate Ab reactivities in experimental autoimmune encephalomyelitis which is the classical rodent model of MS [11,12]. So far, only one peptide microarray study has been conducted with sera from MS patients by

Table 3
Signal intensities of lcSSc-specific peptides.

Peptide sequence	lcSSc_01	lcSSc_02	lcSSc_03	lcSSc_04	lcSSc_05	lcSSc_06	lcSSc_07	lcSSc_08	lcSSc_09	lcSSc_10
MGPRRRSRKPEAPRR (1–15)	**542**	**5819**	**1068**	**1331**	**310**	**15803**	**3534**	**1578**	**16684**	142
TPTPGPSRRGPSLGA (21–35)	22	**6850**	**3628**	321	103	**61949**	**1969**	1324	**2024**	**1055**
GPSRRGPSLGASSHQ (25–39)	43	**2874**	**2119**	125	**564**	**61141**	**4503**	268	**3816**	**810**

Three peptides were filtered as lcSSc-specific by Fisher's exact test. They belong to two N-terminal epitopes of CENPA and share the sequence motif GP-R/S-RR. Numbers in bold indicate that the peptide was filtered as serum-specific according to MAID score and flag information. Signals for the listed peptides were <100 for all control-serum and buffer microarrays.

Quintana et al. [7]. They identified Ab signatures, comprising a panel of CNS protein antigens and heat shock proteins, in clinical and pathologic MS subtypes. As the study was limited to 268 peptides and proteins, there is ample room for further investigations using the high-density peptide array technology.

5. Concluding remarks and perspectives

Peptide microarray data from antibody profiling studies have a sparse and highly skewed distribution. They exhibit other characteristics than DNA microarray data, e.g. huge individual differences and stronger stochastic effects. Data pre-processing and filtering of specific peptides must consider this. We reduced the variability in the fluorescence intensities by averaging over peptide spot triplicates and correction of outliers. The normalization of such data remains an issue as artifacts may be produced if the normalization parameters cannot be estimated reliably. It is important to realize that peptide signals are a measure of Ab affinity and titer, but that other factors have an effect as well, e.g. the individual serum-composition and the specificity (batch) and concentration of the secondary Ab used for staining the microarrays. Secondary Ab bind to the glass surface (requiring a background-correction) and to some of the spotted peptides. In a first filtering step, we therefore discard peptides that present very low signals in all arrays as well as peptides that have a significant signal obtained in the negative control incubation without serum (buffer microarray). An MA plot-based regression curve (MAID filtering) classified peptides as bound by Ab of a particular serum or not. This procedure formed the basis for the second filtering step which for each peptide statistically evaluated the proportion of serum-specific signals observed in the patient versus the control sera. This approach was robust against high signal variations between the patients.

Using this workflow, we could confirm with peptide microarrays that lcSSc patients produce Ab against two antigenic regions of the CENPA protein. On closer inspection, all 10 patients had Ab against either CENPA epitope 1 or 2, but only 6/10 against both. This observation demonstrates that patients can show reactivities against different peptides, which alone are less specific markers for a disease than in combination. Hence, it is important to build multivariate predictors using supervised learning algorithms to achieve better diagnostic and prognostic accuracies with a combined set of peptides. We underlined the need for such multivariate classification approaches in Section 4 regarding the stratification of MS, where already many single auto-Ab targets with moderate specificity have been proposed in the literature. As high-density peptide microarrays allow to measure Ab responses to thousands of potential epitopes simultaneously, they are foreseen to be instrumental for identifying relevant peptide sets.

Additional subsequent bioinformatic analyses can be done at the protein sequence and structure level. If several peptides are found to

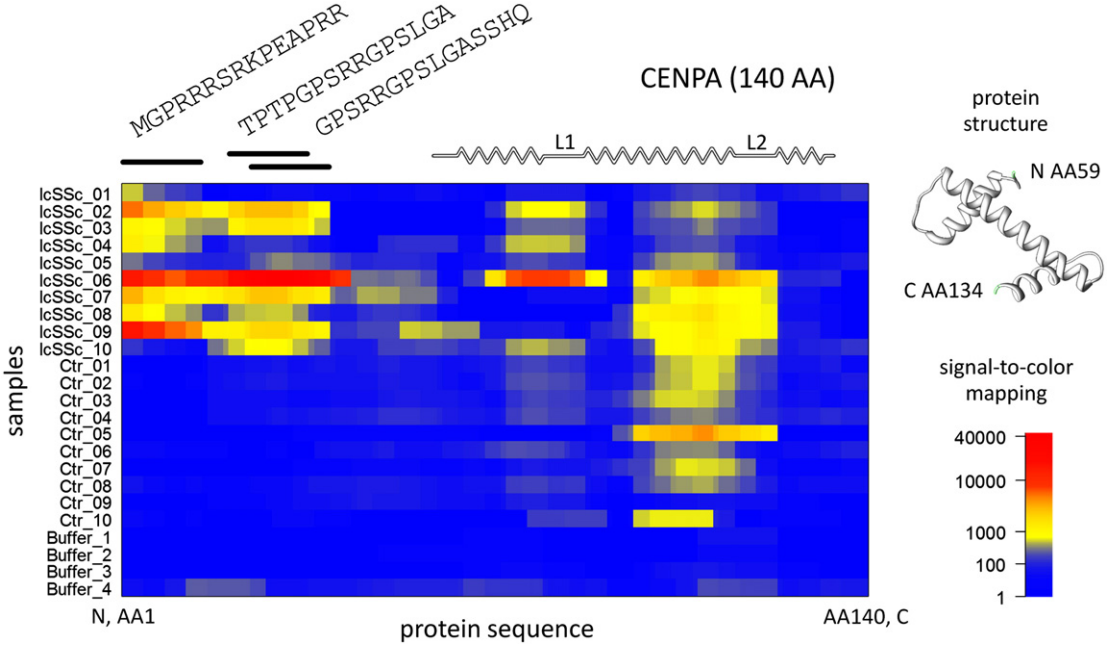

Fig. 5. Antibody reactivities against CENPA peptides. The CENPA protein was represented on the peptide microarrays by 33 peptides. Each row visualizes one microarray incubation. From the left to the right a smoothed signal is shown for each amino acid of the whole protein sequence (the averaged signal of up to five overlapping peptides) (VisualMap). Ab from sera of patients with lcSSc, but not of healthy blood donors, recognize epitopes at the N-terminus of CENPA: antigenic region 1 (1–15) and 2 (21–39). So far, the tertiary structure has only been determined for an N-terminally and C-terminally truncated protein (PDB ID 3NQU [31]). Two regions (69–91 and 97–123) showed elevated signals for patients as well as for healthy controls. These regions encompass the loops between helix α1 and α2 (80–87, L1) and helix α2 and α3 (116–123, L2), respectively.

Table 4
Potential auto-antibody protein targets in multiple sclerosis.

Auto-Ab target	Full name	Sample	Publication	Statement
MBP	Myelin basic protein	Serum	Belogurov et al. [43]	Sera levels of Ab to MBP fragments 48–70 and 85–170 and whole MBP are higher in MS patients than in healthy donors
MOG	Myelin oligodendrocyte glycoprotein	Serum	Wang et al. [44]	MOG Ab in healthy young adults are associated with an increased risk of developing MS
PLP	Proteolipid protein 1	Serum and CSF	Sun et al. [45]	Anti-PLP Ab-secreting cells are detected more frequently in blood and CSF from MS patients than from controls
MAG	Myelin associated glycoprotein	CSF	Baig et al. [46]	12/25 untreated MS patients had cells in CSF secreting anti-MAG Ab, and only 1/27 controls
NEFL	Neurofilament, light polypeptide	CSF	Siber et al. [47]	Patients with progressive MS have elevated NEFL-Ab that may contribute to axonal loss and accumulation of disability
NEFM	Neurofilament, medium polypeptide	CSF	Bartos et al. [48]	Intrathecal Ab against the NEFM subunit are elevated in MS compared with control groups
TUB	Tubulin	CSF	Svarcová et al. [49]	Anti-TUB levels are elevated in MS patients in CSF and synthesized intrathecally
PGAM1	Phosphoglycerate mutase 1	Serum	Kimura et al. [50]	Prevalence of PGAM1 Ab was higher in patients with MS (17/21) and NMO (10/13) than in healthy controls (4/17)
NOGO-A	Reticulon 4	Serum and CSF	Reindl et al. [51]	Ab to Nogo-A are frequent in MS patients, in particular in young patients with relapsing–remitting disease course
NOGOR	Reticulon 4 receptor	Serum	Onoue et al. [52]	None of healthy controls but 33% of MS patients showed anti-non-glycosylated NOGOR Ab
NFASC	Neurofascin	Serum	Mathey et al. [53]	Patients with chronic progressive MS show higher Ab response to neurofascin
CNTN2	Contactin 2	CSF	Derfuss et al. [54]	CNTN2 Ab are more prevalent in MS in comparison to other inflammatory neurological disease
TPO	Thyroid peroxidase	Serum	Annunziata et al. [55]	Anti-TPO detected in 22% of MS patients (5% controls), associated with relapse risk but not with disease duration
CNP	2′,3′-cyclic-nucleotide 3′-phosphodiesterase	Serum and CSF	Lovato et al. [56]	44% of MS patients have serum Ab directed to CNP, which is also frequently recognized by MS CSF (22%)
SPTAN1	Alpha-fodrin	Serum	de Seze et al. [57]	Auto-Ab against the N-terminal portion of alpha-fodrin were detected in 13% of patients with MS
TAL	Transaldolase 1	Serum and CSF	Banki et al. [58]	Ab to TAL were detected in serum (25/87) and CSF (15/20) of patients with MS, and absent in 145 controls
CLDN11	Claudin 11	CSF	Bronstein et al. [59]	Some patients with relapsing–remitting MS contain anti-CLDN11 Ab in their CSF that are not present in controls
TKT	Transketolase	Serum and CSF	Lovato et al. [56]	50% of MS patients CSF Ab (and none of the controls) recognized TKT that is mainly expressed by oligodendrocytes
SPAG16	Sperm associated antigen 16	Serum	Somers et al. [60]	Elevated Ab against SPAG16 isoform 2 were detected in a larger proportion of MS patients compared to controls
CRYAB	Crystallin, alpha B	Serum	Vojdani et al. [61]	Serum anti-alpha-B-crystallin levels are significantly higher in MS patients than in controls
HSP70	Heat stock 70 kDa protein	CSF	Chiba et al. [62]	Higher Ab titers against HSP70 family proteins can be found in CSF from patients with (progressive) MS
HSP90	Heat shock 90 Da protein 1, beta	CSF	Cid et al. [63]	CSF anti-HSP90 Ab levels are significantly higher in MS patients during remission than in control patients
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	CSF	Kolln et al. [64]	GAPDH-Ab were observed more often in MS patients (60%) than in other inflammatory neurological diseases (14%)
hnRNP	Heterogeneous nuclear ribonucleoproteins	CSF	Yukitake et al. [65]	The Ab response to hnRNPs is significantly increased in CSF of MS patients
Proteasome	Proteins of the proteasome complex	Serum	Thuy-Tien et al. [66]	20S proteasome Ab were detected in 13% of relapsing–remitting MS patients (2% controls), reflecting cellular damage

For each protein only one outstanding publication is given as a reference. Ab = antibody, MS = multiple sclerosis, CSF = cerebrospinal fluid, NMO = neuromyelitis optica.

be group-specific, it makes sense to scrutinize them for common sequence patterns (motifs). Filtered peptides with similar sequence or physical–chemical properties might be bound by Ab of the same specificity. Using sequence alignment techniques one can therefore reduce a larger list of peptides to few antigenic sequence motifs. Moreover, one can screen the human proteome (or infectious agents such as viruses) for proteins that match with the filtered peptide sequences. This approach may provide clues about the etiology of autoreactive immune response or identify novel Ab targets. Mapping obtained peptide signal intensities to protein structures is another topic that can be pursued further.

High-density peptide microarrays can be applied in experimental and clinical medicine, for instance to analyze complex Ab repertoires, characterize disease-specific epitope-antibody-recognitions (EAR), and monitor Ab specificities over time (e.g. immune responses to drugs or efficacies of vaccinations). The technology also allows to fine-specify (subtype) Ab reactivities using microarrays displaying mutated epitopes. Moreover, despite the fact that linear epitopes are examined, peptides may also be used to detect Ab that bind conformational epitopes *in vivo* (mimotope recognition). Validation studies with larger cohorts of samples can be done using multi-well microarray solutions (e.g. a 96-well format). However, the true challenge is to wisely choose peptides for measurement. About 3×10^{19} amino acid combinations are possible for peptides with a length of 15 AA. The role of anti-CENPA Ab in SSc shows that not only cell surface proteins are targeted. Therefore, proteins cannot be excluded *per se* for an epitope mapping analysis just because of their cellular localization. One approach to selectively pick peptides might be to take protein structures into account and focus on peptides that are sterically accessible to Ab, and include peptides that could mimic conformational epitopes. On the other hand, during microarray design one might avoid similar peptides and peptides with properties not preferred for Ab-antigen interactions [71]. A panel of peptides optimized in this manner should lead to less sparse data sets and

accelerate our research efforts. Bioinformatic solutions are essential to go this one step further.

Disclosure statement

Prof. Dr. Thiesen is cofounder and Dr. Hecker and Mr. Steinbeck were employees of the Gesellschaft für Individualisierte Medizin mbH (IndyMED, Rostock, Germany), which provides services in performing and analyzing peptide microarray experiments. Dr. Lorenz, Ms. Li Hong, Prof. Dr. Riemekasten, Prof. Dr. Li and Prof. Dr. Zettl declare that they have no competing interest in connection with this paper.

Take-home messages

- Peptide microarrays allow to detect antibody reactivities against thousands of peptides.
- A bioinformatic workflow to analyze peptide microarray data is presented.
- We screened sera of limited cutaneous systemic sclerosis patients for auto-antibodies.
- Patients have antibodies to peptides matching the N-terminus of centromere protein A.
- We reviewed potential diagnostic and prognostic antibody biomarkers in multiple sclerosis.

Acknowledgments

This work was partly financed by regional, national and international grants (EU Autorome LSHM-CT-2004-005264, E-RARE 01GM 0803, MV Excellence UR 08051, BMBF MedSys 0315450D/G, ComBio CHN07/38). We thank Prof. Dr. Mike Schutkowski and Johannes Zerweck from JPT Peptide Technologies GmbH for continued assistance with the peptide microarray technology. Thanks to Dr. Renate Claus and all collaborators in the clinic for providing sera samples as well as to our lab assistant Eva Lorbeer-Rehfeldt for her help in performing the experiments.

References

[1] Robinson WH, Steinman L, Utz PJ. Protein and peptide array analysis of autoimmune disease. Biotechniques Suppl 2002:66–9.
[2] Cretich M, Damin F, Pirri G, Chiari M. Protein and peptide arrays: recent trends and new directions. Biomol Eng 2006;23(2–3):77–88.
[3] Masch A, Zerweck J, Reimer U, Wenschuh H, Schutkowski M. Antibody signatures defined by high-content peptide microarray analysis. Methods Mol Biol 2010;669: 161–72.
[4] Hall DA, Ptacek J, Snyder M. Protein microarray technology. Mech Ageing Dev 2007;128(1):161–7.
[5] Michaud GA, Salcius M, Zhou F, Bangham R, Bonin J, Guo H, et al. Analyzing antibody specificity with whole proteome microarrays. Nat Biotechnol 2003;21 (12):1509–12.
[6] Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, Fries JF, et al. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. Arthritis Rheum 2005;52(9): 2645–55.
[7] Quintana FJ, Farez MF, Viglietta V, Iglesias AH, Merbl Y, Izquierdo G, et al. Antigen microarrays identify unique serum autoantibody signatures in clinical and pathologic subtypes of multiple sclerosis. Proc Natl Acad Sci USA 2008;105(48): 18889–94.
[8] Gaseitsiwe S, Valentini D, Mahdavifar S, Magalhaes I, Hoft DF, Zerweck J, et al. Pattern recognition in pulmonary tuberculosis defined by high content peptide microarray chip analysis representing 61 proteins from M. tuberculosis. PLoS One 2008;3(12):e3840.
[9] Reilly M, Valentini D. Visualisation and pre-processing of peptide microarray data. Methods Mol Biol 2009;570:373–89.
[10] Nahtman T, Jernberg A, Mahdavifar S, Zerweck J, Schutkowski M, Maeurer M, et al. Validation of peptide epitope microarray experiments and extraction of quality data. J Immunol Methods 2007;328(1–2):1–13.
[11] Kidd BA, Ho PP, Sharpe O, Zhao X, Tomooka BH, Kanter JL, et al. Epitope spreading to citrullinated antigens in mouse models of autoimmune arthritis and demyelination. Arthritis Res Ther 2008;10(5):R119.
[12] Robinson WH, Fontoura P, Lee BJ, de Vegvar HE, Tom J, Pedotti R, et al. Protein microarrays guide tolerizing DNA vaccine treatment of autoimmune encephalomyelitis. Nat Biotechnol 2003;21(9):1033–9.
[13] Quintana FJ, Hagedorn PH, Elizur G, Merbl Y, Domany E, Cohen IR. Functional immunomics: microarray analysis of IgG autoantibody repertoires predicts the future response of mice to induced diabetes. Proc Natl Acad Sci USA 2004;101 (Suppl 2):14615–21.
[14] Robinson WH, DiGennaro C, Hueber W, Haab BB, Kamachi M, Dean EJ, et al. Autoantigen microarrays for multiplex characterization of autoantibody responses. Nat Med 2002;8(3):295–301.
[15] Shreffler WG, Beyer K, Chu TH, Burks AW, Sampson HA. Microarray immunoassay: association of clinical history, in vitro IgE function, and heterogeneity of allergenic peanut epitopes. J Allergy Clin Immunol 2004;113(4):776–82.
[16] Neuman de Vegvar HE, Amara RR, Steinman L, Utz PJ, Robinson HL, Robinson WH. Microarray profiling of antibody responses against simian–human immunodeficiency virus: postchallenge convergence of reactivities independent of host histocompatibility type and vaccine regimen. J Virol 2003;77(20):11125–38.
[17] Hachulla E, Launay D. Diagnosis and classification of systemic sclerosis. Clin Rev Allergy Immunol 2011;40(2):78–89.
[18] Hamaguchi Y. Autoantibody profiles in systemic sclerosis: predictive value for clinical evaluation and prognosis. J Dermatol 2010;37(1):42–53.
[19] Mahler M, Fritzler MJ. Epitope specificity and significance in systemic autoimmune diseases. Ann N Y Acad Sci 2010;1183:267–87.
[20] Lennon VA, Kryzer TJ, Pittock SJ, Verkman AS, Hinson SR. IgG marker of optic-spinal multiple sclerosis binds to the aquaporin-4 water channel. J Exp Med 2005;202(4):473–7.
[21] Lindstrom JM, Seybold ME, Lennon VA, Whittingham S, Duane DD. Antibody to acetylcholine receptor in myasthenia gravis. Prevalence, clinical correlates, and diagnostic value. Neurology 1976;26(11):1054–9.
[22] LeRoy EC, Black C, Fleischmajer R, Jablonska S, Krieg T, Medsger Jr TA, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15(2):202–5.
[23] Walker UA, Tyndall A, Czirják L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. Ann Rheum Dis 2007;66(6):754–63.
[24] Lorenz P, Kreutzer M, Zerweck J, Schutkowski M, Thiesen HJ. Probing the epitope signatures of IgG antibodies in human serum from patients with autoimmune disease. Methods Mol Biol 2009;524:247–58.
[25] Hecker M, Goertsches RH, Engelmann R, Thiesen HJ, Guthke R. Integrative modeling of transcriptional regulation in response to antirheumatic therapy. BMC Bioinformatics 2009;10:262.
[26] Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, et al. Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biol 2002;3(11) research0062.
[27] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001;98(9):5116–21.
[28] Mahler M, Mierau R, Blüthner M. Fine-specificity of the anti-CENP-A B-cell autoimmune response. J Mol Med 2000;78(8):460–7.
[29] Muro Y, Azuma N, Onouchi H, Kunimatsu M, Tomita Y, Sasaki M, et al. Autoepitopes on autoantigen centromere protein-A (CENP-A) are restricted to the N-terminal region, which has no homology with histone H3. Clin Exp Immunol 2000;120(1):218–23.
[30] Perosa F, Vicenti C, Racanelli V, Leone P, Valentini G, Dammacco F. The immunodominant epitope of centromere-associated protein A displays homology with the transcription factor forkhead box E3 (FOXE3). Clin Immunol 2010;137 (1):60–73.
[31] Sekulic N, Bassett EA, Rogers DJ, Black BE. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature 2010;467(7313):347–51.
[32] Compston A, Coles A. Multiple sclerosis. Lancet 2008;372(9648):1502–17.
[33] Raine CS, McFarland HF, Hohlfeld R. Multiple sclerosis: a comprehensive text. Philadelphia: Saunders Elsevier; 2008.
[34] Sospedra M, Martin R. Immunology of multiple sclerosis. Annu Rev Immunol 2005;23:683–747.
[35] Lalive PH. Autoantibodies in inflammatory demyelinating diseases of the central nervous system. Swiss Med Wkly 2008;138(47–48):692–707.
[36] Holmøy T. The discovery of oligoclonal bands: a 50-year anniversary. Eur Neurol 2009;62(5):311–5.
[37] Hauser SL, Waubant E, Arnold DL, Vollmer T, Antel J, Fox RJ, et al. HERMES Trial Group, B-cell depletion with rituximab in relapsing–remitting multiple sclerosis. N Engl J Med 2008;358(7):676–88.
[38] Kappos L, Calabresi P, O'Connor P, Bar-Or A, Li D, Barkhof F, et al. Efficacy and safety of ocrelizumab in patients with relapsing–remitting multiple sclerosis: results of a phase II randomized placebo-controlled multicenter trial. Mult Scler 2010;16: S7–S39.
[39] Giovannoni G, Munschauer III FE, Deisenhammer F. Neutralising antibodies to interferon beta during the treatment of multiple sclerosis. J Neurol Neurosurg Psychiatry 2002;73(5):465–9.
[40] Polman CH, O'Connor PW, Havrdova E, Hutchinson M, Kappos L, Miller DH, et al, AFFIRM Investigators. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N Engl J Med 2006;354(9):899–910.
[41] Somers K, Govarts C, Stinissen P, Somers V. Multiplexing approaches for autoantibody profiling in multiple sclerosis. Autoimmun Rev 2009;8(7):573–9.
[42] Eggert M, Zettl UK, Neeck G. Autoantibodies in autoimmune diseases. Curr Pharm Des 2010;16(14):1634–43.
[43] Belogurov Jr AA, Kurkova IN, Friboulet A, Thomas D, Misikov VK, Zakharova MY, et al. Recognition and degradation of myelin basic protein peptides by serum autoantibodies: novel biomarker for multiple sclerosis. J Immunol 2008;180(2): 1258–67.

[44] Wang H, Munger KL, Reindl M, O'Reilly EJ, Levin LI, Berger T, et al. Myelin oligodendrocyte glycoprotein antibodies and multiple sclerosis in healthy young adults. Neurology 2008;71(15):1142–6.
[45] Sun JB, Olsson T, Wang WZ, Xiao BG, Kostulas V, Fredrikson S, et al. Autoreactive T and B cells responding to myelin proteolipid protein in multiple sclerosis and controls. Eur J Immunol 1991;21(6):1461–8.
[46] Baig S, Olsson T, Yu-Ping J, Höjeberg B, Cruz M, Link H. Multiple sclerosis: cells secreting antibodies against myelin-associated glycoprotein are present in cerebrospinal fluid. Scand J Immunol 1991;33(1):73–9.
[47] Silber E, Semra YK, Gregson NA, Sharief MK. Patients with progressive multiple sclerosis have elevated antibodies to neurofilament subunit. Neurology 2002;58 (9):1372–81.
[48] Bartos A, Fialová L, Soukupová J, Kukal J, Malbohan I, Pit'ha J. Elevated intrathecal antibodies against the medium neurofilament subunit in multiple sclerosis. J Neurol 2007;254(1):20–5.
[49] Svarcová J, Fialová L, Bartos A, Steinbachová M, Malbohan I. Cerebrospinal fluid antibodies to tubulin are elevated in the patients with multiple sclerosis. Eur J Neurol 2008;15(11):1173–9.
[50] Kimura A, Sakurai T, Koumura A, Yamada M, Hayashi Y, Tanaka Y, et al. High prevalence of autoantibodies against phosphoglycerate mutase 1 in patients with autoimmune central nervous system diseases. J Neuroimmunol 2010;219(1–2):105–8.
[51] Reindl M, Khantane S, Ehling R, Schanda K, Lutterotti A, Brinkhoff C, et al. Serum and cerebrospinal fluid antibodies to Nogo-A in patients with multiple sclerosis and acute neurological disorders. J Neuroimmunol 2003;145(1–2):139–47.
[52] Onoue H, Satoh JI, Ogawa M, Tabunoki H, Yamamura T. Detection of anti-Nogo receptor autoantibody in the serum of multiple sclerosis and controls. Acta Neurol Scand 2007;115(3):153–60.
[53] Mathey EK, Derfuss T, Storch MK, Williams KR, Hales K, Woolley DR, et al. Neurofascin as a novel target for autoantibody-mediated axonal injury. J Exp Med 2007;204(10):2363–72.
[54] Derfuss T, Parikh K, Velhin S, Braun M, Mathey E, Krumbholz M, et al. Contactin-2/TAG-1-directed autoimmunity is identified in multiple sclerosis patients and mediates gray matter pathology in animals. Proc Natl Acad Sci USA 2009;106(20): 8302–7.
[55] Annunziata P, Lore' F, Venturini E, Morana P, Guarino E, Borghi S, Guazzi GC. Early synthesis and correlation of serum anti-thyroid antibodies with clinical parameters in multiple sclerosis. J Neurol Sci 1999;168(1):32–6.
[56] Lovato L, Cianti R, Gini B, Marconi S, Bianchi L, Armini A, et al. Transketolase and 2′,3′-cyclic-nucleotide 3′-phosphodiesterase type I isoforms are specifically recognized by IgG autoantibodies in multiple sclerosis patients. Mol Cell Proteomics 2008;7(12):2337–49.
[57] de Seze J, Dubucquoi S, Fauchais AL, Matthias T, Devos D, Castelnovo G, et al. Alpha-fodrin autoantibodies in the differential diagnosis of MS and Sjögren syndrome. Neurology 2003;61(2):268–9.
[58] Banki K, Colombo E, Sia F, Halladay D, Mattson DH, Tatum AH, et al. Oligodendrocyte-specific expression and autoantigenicity of transaldolase in multiple sclerosis. J Exp Med 1994;180(5):1649–63.
[59] Bronstein JM, Lallone RL, Seitz RS, Ellison GW, Myers LW. A humoral response to oligodendrocyte-specific protein in MS: a potential molecular mimic. Neurology 1999;53(1):154–61.
[60] Somers K, Hendriks J, Hupperts R, Zwanikken C, Verbeek M, Stinissen P, et al. SPAG16 isoform 2 as a novel autoantigen in multiple sclerosis. J Neuroimmunol 2010;228(1–2):162.
[61] Vojdani A, Vojdani E, Cooper E. Antibodies to myelin basic protein, myelin oligodendrocytes peptides, alpha-beta-crystallin, lymphocyte activation and cytokine production in patients with multiple sclerosis. J Intern Med 2003;254 (4):363–74.
[62] Chiba S, Yokota S, Yonekura K, Tanaka S, Furuyama H, Kubota H, et al. Autoantibodies against HSP70 family proteins were detected in the cerebrospinal fluid from patients with multiple sclerosis. J Neurol Sci 2006;241(1–2):39–43.
[63] Cid C, Regidor I, Alcázar A. Anti-heat shock protein 90beta antibodies are detected in patients with multiple sclerosis during remission. J Neuroimmunol 2007;184 (1–2):223–6.
[64] Kolln J, Ren HM, Da RR, Zhang Y, Spillner E, Olek M, et al. Triosephosphate isomerase- and glyceraldehyde-3-phosphate dehydrogenase-reactive autoantibodies in the cerebrospinal fluid of patients with multiple sclerosis. J Immunol 2006;177(8):5652–8.
[65] Yukitake M, Sueoka E, Sueoka-Aragane N, Sato A, Ohashi H, Yakushiji Y, et al. Significantly increased antibody response to heterogeneous nuclear ribonucleoproteins in cerebrospinal fluid of multiple sclerosis patients but not in patients with human T-lymphotropic virus type I-associated myelopathy/tropical spastic paraparesis. J Neurovirol 2008;14(2):130–5.
[66] Thuy-Tien H, Haugen M, Aarseth J, Storstein A, Vedeler CA. Proteasome antibodies in patients with cancer or multiple sclerosis. Scand J Immunol 2008;67(4):400–3.
[67] Ilyas AA, Chen ZW, Cook SD. Antibodies to sulfatide in cerebrospinal fluid of patients with multiple sclerosis. J Neuroimmunol 2003;139(1–2):76–80.
[68] Schwarz M, Spector L, Gortler M, Weisshaus O, Glass-Marmor L, Karni A, et al. Serum anti-Glc(alpha1,4)Glc(alpha) antibodies as a biomarker for relapsing-remitting multiple sclerosis. J Neurol Sci 2006;244(1–2):59–68.
[69] Jafari N, van Nierop GP, Verjans GM, Osterhaus AD, Middeldorp JM, Hintzen RQ. No evidence for intrathecal IgG synthesis to Epstein Barr virus nuclear antigen-1 in multiple sclerosis. J Clin Virol 2010;49(1):26–31.
[70] McLaughlin KA, Chitnis T, Newcombe J, Franz B, Kennedy J, McArdel S, et al. Age-dependent B cell autoimmunity to a myelin surface antigen in pediatric multiple sclerosis. J Immunol 2009;183(6):4067–76.
[71] Rubinstein ND, Mayrose I, Halperin D, Yekutieli D, Gershoni JM, Pupko T. Computational characterization of B-cell epitopes. Mol Immunol 2008;45(12): 3477–89.

Anti-cytrulinated proteins antibodies and seropositive arthralgia

Anti-cytrulinated protein antibodies (ACPA) may precede the rheumatoid arthritis onset. In an interesting article, van de Stadt et al. (Ann Rheum Dis 2011;70:128-33) have evaluated the role of these antibodies in 244 patients with arthralgia and positibity for ACPA and/or IgM rheumatoid factor. The authors found that 28% of these patients developed arthritis in a median of 11 months. Patients who were positive for at least two ACPA developed arthritis more commonly. In summary, this study demonstrated that in patients with arthalgia and positivity for ACPA, one third of them will evolve with arthritis, maninly those with positivity for at least to antibodies.

Jozélio Freire de Carvalho MD, PhD