Bioorganic & Medicinal Chemistry 19 (2011) 2470–2477

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

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

Assessing high affinity binding to HLA-DQ2.5 by a novel peptide library based approach

Ulrike Jüse [a,*], Magnus Arntzen [b,c,d], Peter Højrup [e], Burkhard Fleckenstein [f], Ludvig M. Sollid [a,f]

[a] Centre for Immune Regulation, Institute of Immunology, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway
[b] Proteomics Core Facility, Oslo University Hospital, Rikshospitalet & University of Oslo, 0372 Oslo, Norway
[c] The Biotechnology Centre of Oslo, University of Oslo, 0317 Oslo, Norway
[d] Proteomics Core Facility, Norwegian University of Life Sciences, 1432 Ås, Norway
[e] Institute of Biochemistry and Molecular Biology, University of Southern Denmark, 5230 Odense, Denmark
[f] Centre for Immune Regulation, Institute of Immunology, University of Oslo, 0372 Oslo, Norway

ARTICLE INFO

Article history:
Received 13 March 2010
Revised 23 January 2011
Accepted 27 January 2011
Available online 2 February 2011

Keywords:
Dedicated peptide libraries
HLA-DQ2.5
Mass spectrometry
Peptide blockers
Celiac disease

ABSTRACT

Here we report on a novel peptide library based method for HLA class II binding motif identification. The approach is based on water soluble HLA class II molecules and soluble dedicated peptide libraries. A high number of different synthetic peptides are competing to interact with a limited amount of HLA molecules, giving a selective force in the binding. The peptide libraries can be designed so that the sequence length, the alignment of binding registers, the numbers and composition of random positions are controlled, and also modified amino acids can be included. Selected library peptides bound to HLA are then isolated by size exclusion chromatography and sequenced by tandem mass spectrometry online coupled to liquid chromatography. The MS/MS data are subsequently searched against a library defined database using a search engine such as Mascot, followed by manual inspection of the results. We used two dodecamer and two decamer peptide libraries and HLA-DQ2.5 to test possibilities and limits of this method. The selected sequences which we identified in the fraction eluted from HLA-DQ2.5 showed a higher average of their predicted binding affinity values compared to the original peptide library. The eluted sequences fit very well with the previously described HLA-DQ2.5 peptide binding motif. This novel method, limited by library complexity and sensitivity of mass spectrometry, allows the analysis of several thousand synthetic sequences concomitantly in a simple water soluble format.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

HLA associations are seen for most autoimmune diseases. HLA-DQ2.5 (DQA1*0501/DQB1*0201) of the B8-DR3-DQ2 haplotype is associated with particularly many diseases, including type 1 diabetes, Addison's disease, systemic lupus erythematosus, and celiac disease.[1–3] Celiac disease is a complex inflammatory disorder of the small intestine, and HLA-DQ2.5 exerts a central role in the pathogenesis of this disease by presenting gluten peptides to specific CD4+T-cells in the intestinal mucosa.[4] Over 90% of the celiac disease patients express HLA-DQ2.5.[5,6] For celiac disease, but also other HLA-DQ2.5 associated diseases, specific interference with this key molecule can represent a novel treatment strategy. An HLA blocker should interfere with the binding of immunogenic peptide(s), but should not be recognized by any T cell receptor. So far, the most promising approaches for blocking of the HLA-DQ2.5 mediated presentation of gluten derived antigens is the use of cyclic and dimeric gluten peptides,[7] or the introduction of a large, chemically modified side chain into gluten epitopes.[8] However, the efficacy of the designed blockers for inhibiting T cell activation in vitro is inadequate, and it appears that insufficient affinity of the blockers for HLA-DQ2.5 is at least one limitation.

In this study we present a method to identify optimal HLA class II ligands. By incubating peptide libraries with several randomized positions with limited amounts of HLA molecules, an environment of strong competition is created in which HLA-DQ2.5 selects for the best binders (Fig. 1). Identification of the selected ligands is done by mass spectrometry, and thus also modified amino acids can be included in the libraries. We here report on the strengths and weaknesses of the method and present data on HLA-DQ2.5 ligands and the peptide binding motif of this molecule.

2. Results

2.1. Peptide libraries with a single randomized position—libraries A and B

Peptide libraries with a single randomized position were used to optimize separation of HLA-DQ2.5-peptide complexes from

* Corresponding author. Tel.: +47 23 07 30 17; fax: +47 23 07 35 10.
E-mail address: ulrike.juse@rr-research.no (U. Jüse).

0968-0896/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2011.01.057

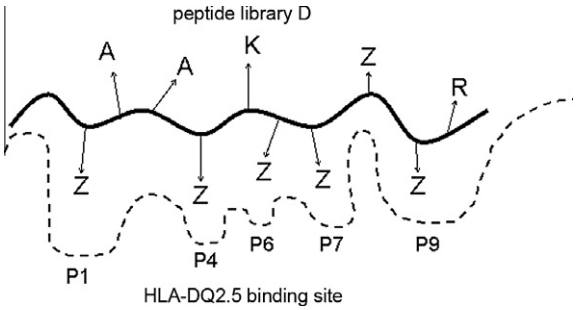

Figure 1. Schematic representation of peptide library D registering into the peptide binding groove of HLA-DQ2.5. Five of the six dedicated Z positions interact with the five pockets in the groove and one is bridging a ridge in position P8. The positively charged amino acid lysine is placed in the non-anchor position P5.

unbound peptides and to explore the possibilities of this novel method. Due to the low complexity, the libraries could be easily analyzed by MALDI-TOF mass spectrometry without the need for LC-based peptide separation. In addition, as each peptide sequence is defined by its $[M+H]^+$ value, sequence identification of the detected peptides by database searching was not required. Notably, for some peptides the isotopic envelopes overlap and the contribution of the different peptides to signal intensity was calculated by using their theoretical isotopic envelope (Supplementary Fig. S1).

We used two dodecapeptide libraries each carrying a single X position in the anchor positions P1 or P9 (Table 1).

The mass spectra of library A, which were acquired from the original peptide library and from the peptides eluted from HLA-DQ2.5, are shown in Fig. 2. All $[M+H]^+$ values calculated from the format of this library were observed in the mass spectrum of the original library. For the eluted sample, significant signals (18 > x for Y) were only obtained for $[M+H]^+$ values which corresponded to sequences with bulky hydrophobic amino acids in position P9 (V, L/I, F, Y, W). Library B, carrying the randomized position P1, was analyzed by the same method. For the original library, $[M+H]^+$ values corresponding to all expected peptide sequences, except the one with proline in the X position, were identified. By contrast to library A, the effect of the selection by HLA-DQ2.5 was less pronounced for library B as we found signals for all 19 peptide sequences in the eluted sample (data not shown).

2.2. Peptide library with four dedicated Z positions—library C

LC–MS/MS analysis of peptide library C on an LTQ-Orbitrap XL mass spectrometer resulted in 5059 MS/MS events for the original library (Table 2). A Mascot search of these MS/MS events against a library defined database led to the identification of 669 peptides (Table 2). The RD values for each of the amino acids in the dedicated Z positions of these peptides were in the range of 0.5–1.5, not far from the expected RD value of 1 (Fig. 3A).

Table 1
Peptide libraries used in this study

		Peptide binding register										Sequences
		1	2	3	4	5	6	7	8	9		
Library A	GG	F	A	P	E	K	E	E	P	**X**	R-NH_2	19
Library B	GG	**X**	A	P	E	K	E	E	P	F	R-NH_2	19
Library C	Ac	A	A	A	A	K	**Z**	**Z**	**Z**	**Z**	R-NH_2	1430
Library D	Ac	**Z**	A	A	**Z**	K	**Z**	**Z**	**Z**	**Z**	R-NH_2	64,350

Libraries A and B carry each a single and fully randomized X position which is composed of an equimolar mixture of 19 proteinogenic amino acids (excluding cysteine). Libraries C and D represent dedicated peptide libraries with four and six dedicated Z positions whose composition are given below.
Amino acid composition in the Z positions of dedicated libraries C and D.
Z in position 1: P, V, E, W, F.
Z in position 4: P, V, T, G, W, L, Y, N, E.
Z in position 6: Q, T, N, Y, F, V, G, L, A, P, E.
Z in position 7: G, L, F, A, E.
Z in position 8: P, E.
Z in position 9: G, Q, A, S, V, P, L, E, Y, D, M, F, W.

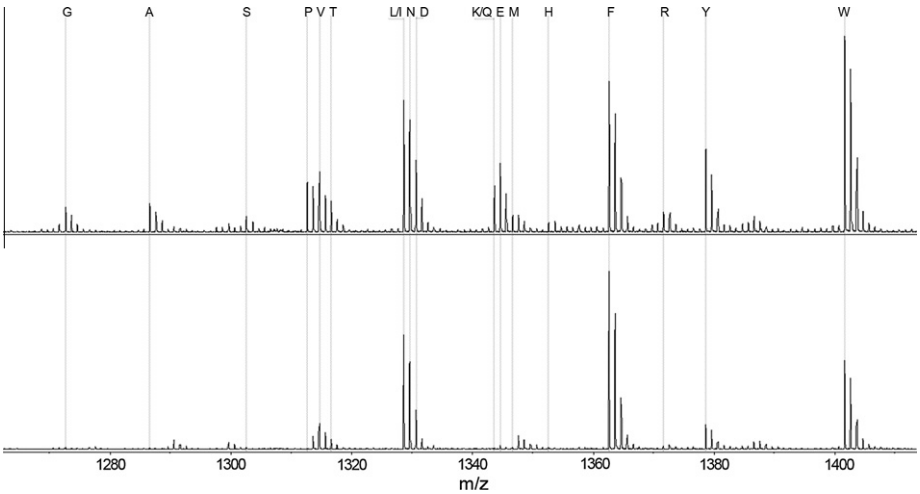

Figure 2. MALDI-TOF mass spectrum of the original library A (top) and of the sample obtained after incubation of HLA-DQ2.5 with library A and peptide elution (bottom). Signals are assigned by the amino acid in the randomized position P9 of the identified peptide sequence.

Table 2
Number of MS/MS events and identified peptides in experiments with libraries C and D

	Library C	Library D
Theoretical sequences	1430	64,350
MS/MS events acquired before incubation	5059	5791
MS/MS events acquired in the eluted sample	707	1212
Sequences identified before incubation	669	1985
Sequences identified in the eluted sample	81	261

LC–MS/MS analysis of the sample after elution from HLA-DQ2.5 resulted in 707 MS/MS events (Table 2). Mascot reported 81 unique top ranking sequences, however, for the majority of these 81 MS/MS events, more than one peptide sequence was assigned by a significant score. Therefore, all of these MS/MS events were subjected to additional manual data inspection, including examination of their retention times and elution profiles, as isobaric peptides can differ slightly in their LC retention times (Fig. 4). Additionally, for these 81 MS/MS events, we manually compared the observed fragment ions with their theoretical fragmentation patterns. This manual data evaluation resulted in the exclusion of 12 sequences and inclusion of 12 sequences. Thus, 81 unique sequences were identified from library C after selection by HLA-DQ2.5 (Supplementary Table S1).

Based on these final 81 identified sequences, RD values were calculated for each amino acid present in the four dedicated Z positions. These RD values indicated that HLA-DQ2.5 favored valine, alanine, and glutamate at position P6, glutamate in position P7 and P8 and valine, leucine, glutamate, tyrosine, aspartate, methionine, phenylalanine, and tryptophan at position P9 (Fig. 3B). In contrast, a few amino acids were not detected or present in low amounts in the dedicated Z positions. These were glutamine, threonine, asparagine, and glycine at position P6, glycine at position P7, and glycine, serine, and proline at position P9 (Fig. 3B).

For all peptides theoretically present in the original library C, and for the 81 peptides identified in the eluted sample, we calculated the sum of the *relC* values ($\sum relC$). For the eluted peptides the average of these $\sum relC$ values was increased by 1.7 compared to the average of the $\sum relC$ values of all peptides of library C (Fig. 5). This finding suggests that HLA-DQ2.5 has indeed selected for optimal HLA-DQ2.5 ligands.

2.3. Peptide library with six dedicated Z positions—library D

LC–MS/MS analysis of peptide library D resulted in 5791 MS/MS events for the original library and 1212 MS/MS events for the eluted peptide sample after incubation with HLA-DQ2.5 (Table 2). The MS/MS data obtained for the original library were subjected to a Mascot search, and we could identify 1985 of the 65,350 theoretically possible peptide sequences (Table 2). Based on these peptides, we calculated RD values for the amino acids in the dedicated Z positions, and found that they varied between 0.1 and 3.8 (Fig. 6A).

After peptide selection by HLA-DQ2.5, the obtained MS/MS data were searched by Mascot resulting in the identification of 179 unique top ranking peptide sequences. All MS/MS events that had more than one peptide sequence assigned by a significant score were manually inspected. As only 3% of the theoretical sequences of library D were identified in the sample before incubation with HLA-DQ2.5, this manual evaluation did not include comparison of retention times and elution profiles but consisted only of comparison of the observed with the theoretical fragmentation patterns. This manual data evaluation increased the number of identified sequences to 261 (Table 2, Supplementary Table S2). The average of the $\sum relC$ values of the eluted peptides was increased by 3.3 compared to the average of the $\sum relC$ values of all peptides present in the original library D (Fig. 5). This suggests an enrichment of optimal HLA-DQ2.5 binding peptides from library D also.

As RD values of amino acids in the dedicated positions of the original library D showed substantial deviation from 1, there is a likely bias in the RD values calculated for the HLA-DQ2.5 selected sample (Fig. 6B). In order to compensate for this bias, we calculated adjusted RD values by dividing the RD values obtained for the HLA-DQ2.5 selected sample by the RD values obtained for the original library (Fig. 6C). These adjusted RD values suggest that HLA-DQ2.5

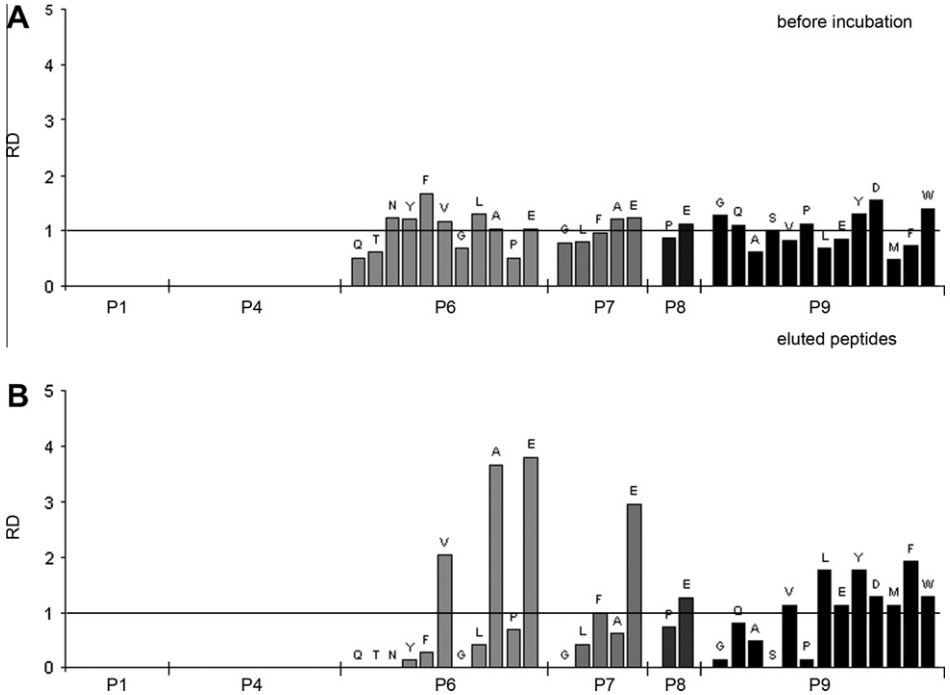

Figure 3. Relative distribution (RD) of amino acids present in the dedicated positions of peptide library C before incubation (A) and after elution (B).

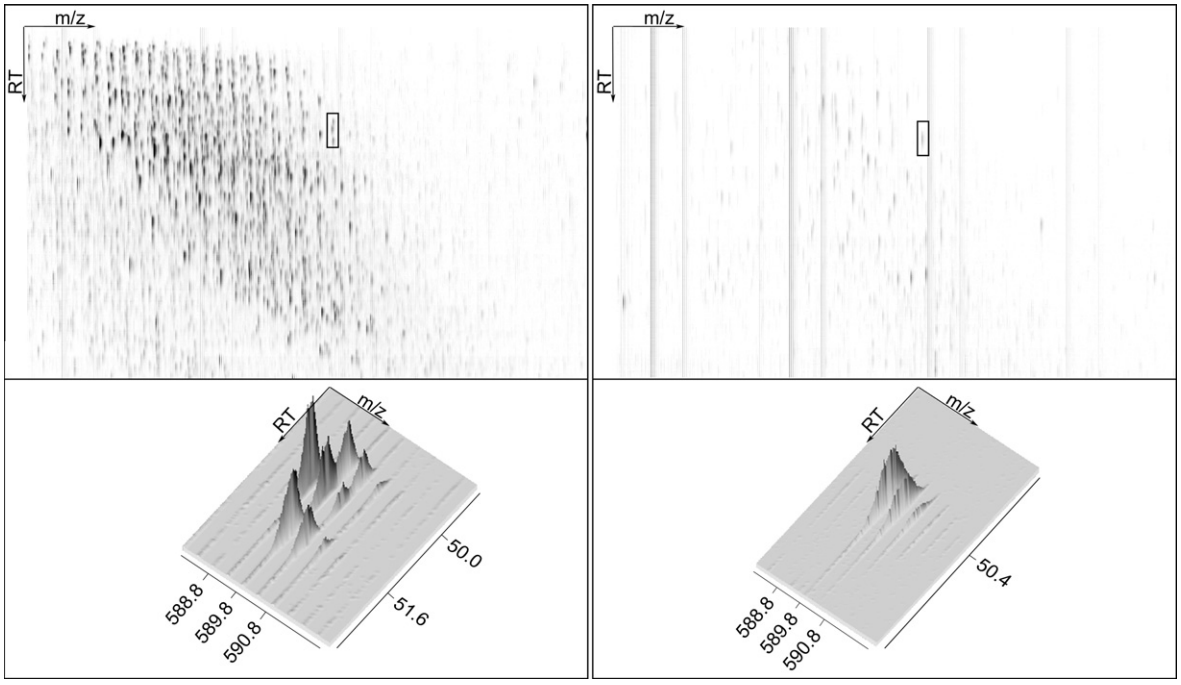

Figure 4. LC–MS map of library C before incubation with HLA-DQ2.5 (left panels) and after elution from HLA-DQ2.5 (right panels). The upper panels show the elution profiles in the region *m*/*z* 425–725 and with retention times from 41 to 73 min. The lower panels present an enlarged 3D view of the rectangular boxes included in the upper panels. Before incubation with HLA-DQ2.5, the lower panel shows three peptides with slightly different retention times (observed as doubly charged ions). The peptides eluting at 50.1 and 51.6 min are isobaric with *m*/*z* 589.9 and MS/MS data allowed identifying their sequences as AAAAKEEEYR and AAAAKYEEER, respectively. After incubation with HLA-DQ2.5 only one peptide was observed at an aligned retention time of 50.4 min which was identified as AAAAKEEEYR. Images were created with MSight version 2.A.6 and aligned retention times were calculated using POSTMan.

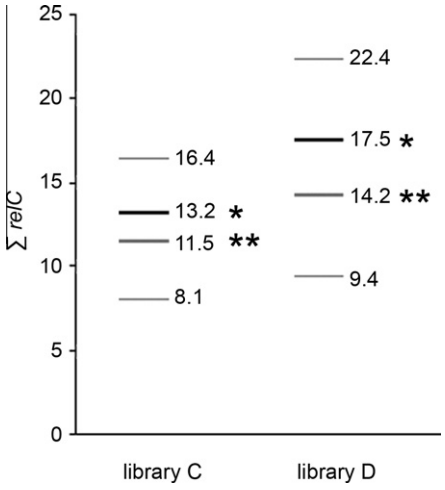

Figure 5. The range of ∑*relC* values for all peptides present in libraries C and D. The average of ∑*relC* values of the original library (**) and of the peptides eluted from HLA-DQ2.5 (*) are given.

favored phenylalanine and tryptophane at position P1; valine, glutamate, and leucine at position P4; valine, glutamate, proline, alanine, and leucine at position P6; glutamate and leucine at position P7, glutamate at position P8, and phenylalanine, leucine, tryptophan, tyrosine, and methionine at position P9 (Fig. 6C). Several amino acids were not detectable or present at low abundance at some positions, when the peptide sample selected by HLA-DQ2.5 was analyzed. These included proline, valine, and glutamate at position P1; threonine, glycine, tryptophane, and tyrosine at position P4, glutamine at position P6; glycine at P7; and glycine, glutamine, serine, valine, and glutamate at position P9.

2.4. HLA-DQ2.5-peptide binding of three peptides identified from library D

In peptide binding assays we analyzed the HLA-DQ2.5 binding ability of three random peptides (FAALKEEEYR, FAAEKVEEFR, FAAEKEEELR) with high scores from library D (Fig. 8). The binding affinity of these peptides was compared to the binding affinity of the high affinity peptides P198 (MB 65 kDa 243-255Y, KPLLIIAEDVEGEY) and P418 (MHC I α 46-60, EPRAPWIEQEGPEYW) and the DQ2-α-I-gliadin epitope (QLQPFPQPELPY). The three peptides derived from library D were found to bind to HLA-DQ2.5 with high affinity. FAAEKEEELR bound with highest affinity with a relative binding capacity (RBC) of 0.93 compared to the high affinity reference peptide P198.

3. Discussion and conclusions

Here we report on a new method for the specific enrichment of peptides with high affinity to a HLA class II molecule from complex soluble dedicated peptide libraries. This method allows identification of high affinity ligands as well as refined definition of peptide binding motifs.

Interfering with antigen presentation by blocking the binding groove of HLA class II molecules is a conceivable strategy for the treatment of HLA associated diseases. HLA-DQ2.5 is a particular attractive target, as many diseases show associations with this HLA molecule.[9] Development of HLA-DQ2.5 specific blockers has however so far given limited results, and an obvious limitation is the HLA binding affinity of the blockers. Thus there is a need for tools that allow the identification of high affinity ligands and that allow for testing of compounds incorporating modified amino acids. Here we present a peptide library based method, that should work equally well with natural as well as modified amino acids, as

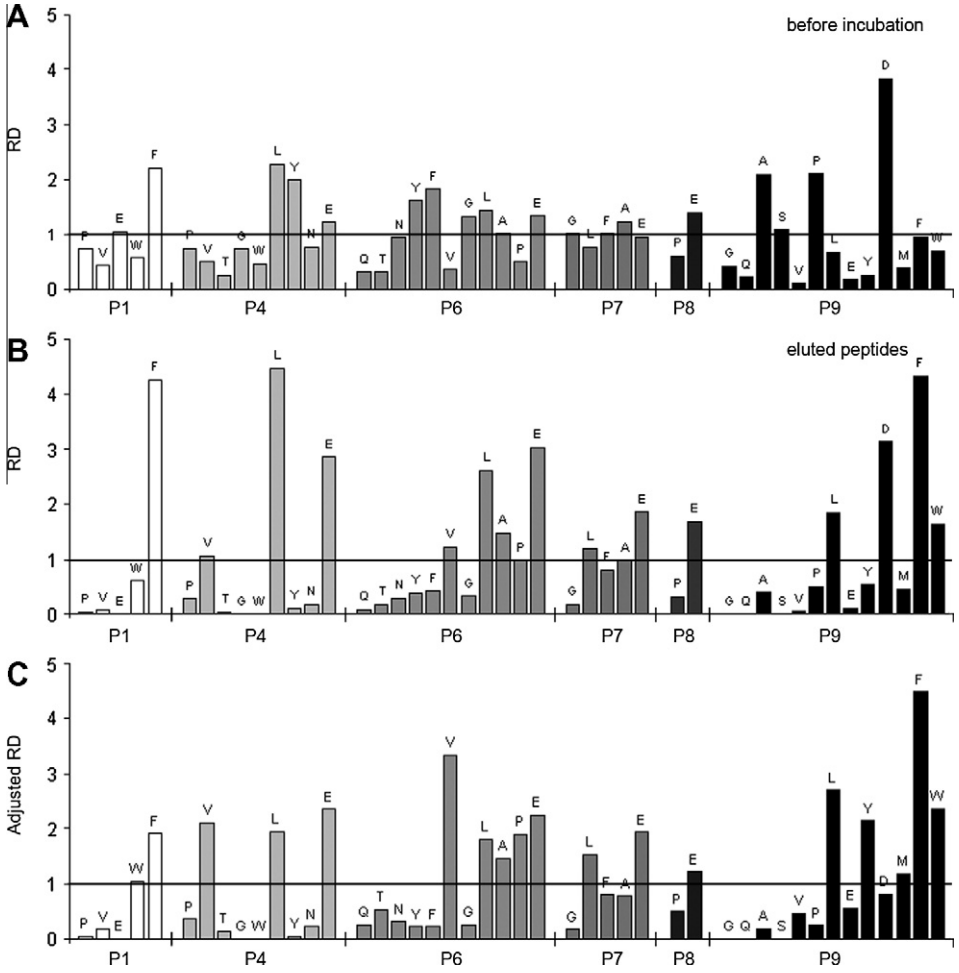

Figure 6. Relative distribution (RD) of amino acids present in the dedicated positions of peptide library D before incubation (A) and after elution (B). In order to compensate for a potential bias in these data, RD values given in B were divided by RD values given in A resulting in adjusted RD values (C).

the detection is undertaken by mass spectrometry. The proof of principle of the method has been established by use of natural amino acids and dedicated peptide libraries.

We first analyzed peptide libraries with a single fully randomized X position in the anchor positions P1 and P9. By comparing the MALDI-TOF mass spectra of the peptide library before incubation with the spectra after elution, we found a specific depletion of several peptides. Peptides with bulky hydrophobic amino acids at position P9 were favored by HLA-DQ2.5 and were the only sequences observed in the eluted sample. At position P1 however, HLA-DQ2.5 was less selective and bound all of the 19 tested peptides. Prior studies have shown, that HLA-DQ2.5 both at positions P9 and P1 prefer amino acids with bulky hydrophobic side chains.[10,11] However, many immunogenic gliadin peptides have the non-bulky amino acid proline at position P1, but never in position P9.[12–14] This suggests that at position P1, a non-preferred amino acid can be selected, provided all other anchors are optimal.

Next we designed two dedicated peptide libraries with four and six dedicated positions (Z). For peptide library C, with four dedicated Z positions, we chose a decapeptide library format with four neighboring dedicated Z positions in the C-terminal part. Notably, the N-terminal anchor positions P1 and P4 were occupied by alanine, a non-optimal anchor residue for HLA-DQ2.5. By using this format, we sought to solely identify sequences with optimal residues in all of the three anchor positions P6, P7, and P9. Each of the dedicated Z positions was composed of 2–13 amino acids, that previously have been demonstrated to be intermediate or good residues in the corresponding positions (Fig. 7).[15] In peptide library D, two additional dedicated Z positions were placed in the two remaining anchor positions, P1 and P4. These Z positions were composed of altogether five (at P1) and nine (at P9) amino acids, that previously have been demonstrated to be intermediate and good anchor residues for these positions (Fig. 7). The two additional Z positions increased the complexity of the peptide library by 50-fold compared to library C. In library D, a total of 65,000 sequences competed for binding to HLA-DQ2.5. Due to the complexity it was necessary to separate peptides in libraries C and D by liquid chromatography prior to MS/MS analysis, and for sequence identification a library defined database and a search engine (Mascot) was required.

The HLA-DQ2.5 selected sequences identified from both the C and D libraries had higher average $\sum relC$ values than the peptides each library was composed of. This is suggesting that the HLA-DQ2.5 affinity matrix did select for ligands with increased affinity in both cases. From library C we identified a panel of 81 peptide ligands (Table 2, Supplementary Table S1), and we established RD values which designated preferred and less preferred anchor residues at the P6, P7, and P9 positions (Fig. 3B). Notably, about half of the theoretically possible sequences were identified when analyzing the original library C. The approximately equal distribution of amino acids in the dedicated positions, argues for a high quality of the peptide libraries used in this study.

The data obtained from library D is less informative. There was an uneven representation of amino acids in the dedicated positions

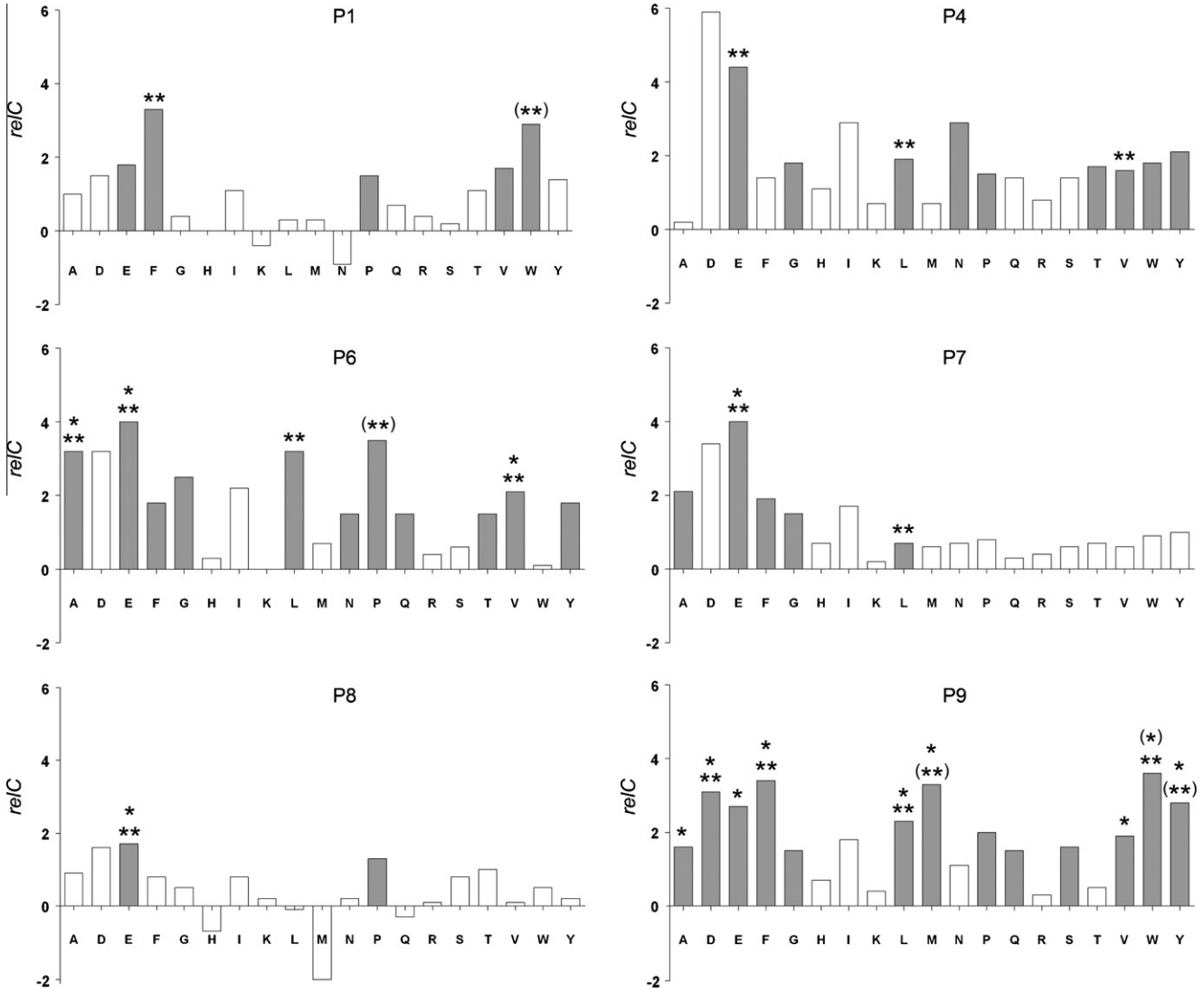

Figure 7. The amino acids in the dedicated positions P1, P4, P6, P7, P8, and P9 in libraries C and D were chosen based on their relative competition values (*relC*).[15] Unfavorable amino acids have low *relC* values; favorable amino acids have high *relC* values. Residues included in dedicated positions of libraries C and D are indicated in gray. In each dedicated position, mainly intermediate and good residues in the corresponding position with a *relC* value greater 1.5 were included (except L in position P7 and P in position 8). The most abundant residues identified in the peptides eluted from HLA-DQ2.5 after incubation are marked with stars (library C = *, library D = **). By comparing RD values of the eluted sample of library D with the RD values of the sample before incubation, some additional amino acids are more abundant in the eluted sample then before incubation. These amino acids are labeled with (**). In library C, tryptophan (W) in position P9 is less abundant in the eluted sample then in the sample before incubation (*).

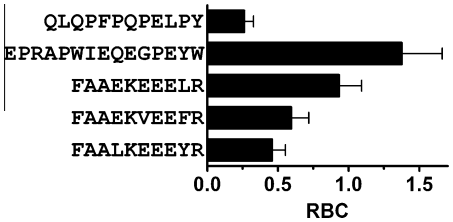

Figure 8. The relative binding capacity (RBC) of three example peptides identified from library D for binding to HLA-DQ2.5. The IC_{50} values of these N-terminally acetylated and C-terminally amidated decapeptides were compared to the IC_{50} value of the reference peptide P198 (MB 65 kDa 243-255Y, KPLLIAEDVEGEY; IC_{50} = 2.3 μM). The high affinity binder P418 (MHC I α 46–60, EPRAPWIEQEGPEYW) and the α-I-gliadin epitope (QLQPFPQPELPY) were also included in the experiment. The values (mean and standard deviation) were obtained from three independent fourfold titration experiments.

of the original library which was probably introduced during the synthesis of the library, or which is due to a random error, introduced when Mascot is searching a database of that high complexity. The list of identified peptides from the sample selected by HLA-DQ2.5 might therefore also be biased. We therefore adjusted the RD values for these selected peptides, and these adjusted RD values give some insight into the HLA-DQ2.5 binding motif.

Further, we analyzed the binding to HLA-DQ2.5 of three randomly selected peptides with high scores from library D in a competitive peptide binding assay (Fig. 8). All three peptides were found to bind to HLA-DQ2.5 with high affinity. The binding affinity of the strongest binder (FAAEKEEELR) was comparable to the binding affinity of the high affinity reference peptide P198.

Our approach can easily be extended to the incorporation of modified amino acids into the peptide libraries, as the identification of selected ligands is done by tandem mass spectrometry. It appears that efficient HLA-DQ2.5 blockers can not solely be made of natural amino acids.[7] Ongoing research has started exploring high affinity binders to HLA-DQ2.5 with both proteinogenic and modified, non-natural amino acids.[7,8] Hopefully this work can lead to the development of blockers which show sufficient efficacy in vitro and in vivo.

4. Experimental

4.1. Soluble HLA-DQ2.5 molecules, expression and purification

Water soluble HLA-DQ2.5 molecules (DQA1*0501/DQB1*0201) tethered with the DQ2-α-I-gliadin epitope (sDQ2-αI) were expressed and purified as described before.[16] The purity of HLA-DQ2.5 preparations was analyzed by SDS–PAGE. Protein concentrations were determined by a BCA protein assay kit (Pierce, Rockford, IL, USA).

4.2. Peptide library design

We tested altogether four peptide libraries based on the same sequence scaffold (Table 1). Two dodecapeptide libraries, named A and B, were randomized at one position each (position X) which was composed of an equimolar mixture of all natural amino acids, except cysteine. These X positions were either placed at anchor position P9 (library A) or anchor position P1 (library B) of the HLA-DQ2.5 binding register (Table 1). Further, we used two acetylated decapeptide amide libraries with four and six dedicated positions with limited variation. In these libraries the dedicated positions (positions Z) were composed of an equimolar mixture of 2–13 different proteinogenic amino acids and placed in several anchor positions and one non-anchor position of the HLA-DQ2.5 binding frame (Fig. 1, Table 1). In library C, the Z positions were placed in positions P6, P7, P8, and P9, and in library D, the Z positions were localized in positions P1, P4, P6, P7, P8, and P9. The amino acids selected for the Z positions of the libraries C and D were residues assumed to be intermediate to optimal anchors based on their relative competition (*relC*) values (>1.5) as previously established by a scanning peptide library approach (Fig. 7).[15] Further, both libraries carried lysine in position P5 and arginine in position P10 to prevent binding of each peptide in multiple binding registers. In all other positions alanine was used.

4.3. Peptide library synthesis

Peptide libraries with one X position were synthesized on Rink amide MBHA-resin using Fmoc/HBTU chemistry and a pipetting robot (Syro I, MultiSynTech, Bochum, Germany). The randomized X position was coupled by using an equimolar mixture of all proteinogenic amino acids except cysteine. Coupling was performed twice using equimolar amounts of amino acids and available amino groups. Libraries C and D were purchased from EZBiolab (EZBiolab Inc., Carmel, USA).

4.4. Competitive binding conditions

sDQ2-αI molecules were treated with thrombin (0.1 U per 4.6 μg of sDQ2-αI) for 2 h at room temperature, followed by a treatment with 0.5 μg 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride. Subsequent, 50 μg peptide library and 1.5 μg soluble HLA-DM (a gift from E. Mellins) were added to 8.3 μg thrombin-cleaved sDQ2-αI in a total volume of 80 μl phosphate buffered saline containing 22.5 mM citric acid and 50.5 mM Na_2HPO_4 (pH 5.3). The sample was incubated for 48 h at 37 °C.

4.5. Separation of high affinity peptides by size exclusion chromatography

Library derived peptides bound to sDQ2 were separated from unbound peptides by preparative size exclusion chromatography on an Äkta purifier system (Amersham Bioscience Corp., Piscataway, NJ, USA) using a Superdex 200 10/300 GL column (Amersham Bioscience) and 15 mM phosphate buffered saline as the mobile phase. The sDQ2 containing fraction was concentrated and sDQ2-bound peptides were eluted by adjusting to pH 1 with trifluoroacetic acid. The samples were incubated for 30 min and peptides were purified on a reversed phase (Poros 20 R2, Applied Biosystems, Foster City, CA, USA) packed in 20-μl GELoader tips (Eppendorf, Hamburg, Germany). Peptides were eluted with 10 μl 99.9% acetonitrile and the sample was dried.

4.6. Identification of peptides from libraries with a single X position by MALDI-TOF mass spectrometry

A volume of 1 μl of the eluted peptide samples was applied on a stainless steel MALDI target plate and 0.5 μl of 10 g/L α-cyano-4-hydroxycinnamic acid in 70% acetonitrile/0.1% TFA was added. MALDI-TOF mass spectra were acquired on an Ultraflex II instrument (Bruker Daltonics, Bremen, Germany) and peptide sequences were identified manually. For all peptides with an overlapping isotopic envelope, the theoretical isotopic distribution was taken into consideration.

4.7. LC–MS/MS analysis of samples from dedicated libraries

The dried peptide samples were dissolved in 0.1% formic acid and analyzed by nanoLC–MS/MS using an EASY nLC system (Proxeon A/S, Odense, Denmark) coupled to a LTQ-Orbitrap XL mass spectrometer (ThermoScientific, Bremen, Germany) equipped with a nano-electrospray ion source. The peptides were separated on a 100 μm ID capillary column 18 cm length, packed with Reprosil C_{18}, 3 μm. The instrument was set to a linear gradient of 0–34% B for 50 min (solvent B was 95% acetonitrile), the threshold for fragmentation was set to 15.000 and the range was 300–1800. Of the original library samples, 0.4 μg were analyzed by using the same settings.

4.8. Peptide identification after analysis of samples from dedicated libraries

Obtained LC–MS/MS data were extracted from the Orbitrap raw files using the Proteome Discoverer software (Thermo Fisher Scientific Inc., Waltham, MA, USA) with default settings and searched against a library defined database using Proteome Discoverer and the Mascot search engine (Matrixscience Ltd, Boston, MA, USA). The following settings were selected: amidated C-terminus and acetylated N-terminus as fixed modifications, Arg-C as the enzyme, no missed cleavages, peptide mass tolerance of 10 ppm, and fragment ion mass tolerance of 0.6 Da. Search results for the original library and for the peptide sample selected by HLA-DQ2.5 were copied to an Excel spreadsheet giving information on the peptide sequences reported for each MS/MS event, together with their ion scores, rank order and retention times. Due to the format of the library, Mascot assigned several isobaric peptides with high scores for most of the MS/MS events. As these isobaric peptides share many of their fragment ion masses, the exact sequence of a selected parent ion can often only be differentiated on the basis of one or a few y- and/or b-fragments. If more than one hit was reported by Mascot for a given MS/MS event in the eluted samples, manual data inspection was performed. For both dedicated libraries, the theoretical fragmentation pattern of all hits reported by Mascot for a particular MS/MS event were compared to the fragments observed in the acquired MS/MS spectrum. For library C, the retention times and elution profiles for all isobaric peptide sequences reported for an MS/MS event were additionally compared with the retention times and elution profiles of all of the isobaric peptides in the sample of the original peptide library. For that reason, retention times for the analysis of the original and selected library C were aligned using the software POSTMAN.[17] By this manual

data inspection, several top ranking hits were excluded as false positive hits whereas other hits were included which originally were not assigned by top scores in the Mascot search.

4.9. Analysis of amino acid preference at given positions

Relative distribution (RD) values, given as the frequency of an amino acid at a dedicated Z position divided by the total number of amino acids represented at the same position, were calculated to visualize amino acid preference. For example phenylalanine (F) in library C was identified 12 times in position P9 in the sample after incubation. A total of 81 peptides were identified and position P9 was composed of 13 amino acids. This gives an RD value of 1.93 (12/81*13). At a given position, favored amino acids have RD values larger than one, and non-favored amino acids have RD values smaller than 1.

4.10. Peptide binding and competition assay

Three example peptides identified from library D (FAALKEEEYR, FAAEKVEEFR, FAAEKEEELR) were purchased as amidated and acetylated peptides (Schafer-N (Copenhagen, Denmark) and tested for binding to HLA-DQ2.5. The competition binding assay was mainly performed as described[18] using the Epstein-Barr virus (EBV)-transformed B-cell line VAVY (IHWS #9023) as source of HLA-DQ2.5 molecules (lysate of 2×10^5 cells/well), the HLA-DQ-specific mAb SPV-L3 (2 μg/well) as catching antibody and the biotinylated peptide P418 (0.5 μM) as indicator peptide. The inhibitory effect on binding of the indicator peptide by the competitor peptides was titrated, and IC_{50} (half maximal inhibitory concentration) values were established for each peptide. The relative binding capacity (RBC) was calculated as the ratio of the IC_{50} value of the reference peptide P198 (MB 65 kDa 243-255Y, KPLLIIAEDVEGEY) to the IC_{50} value of the tested peptide. Three independent fourfold titration experiments were performed.

Acknowledgments

The financial support from the European Commission (Marie Curie Research Training Network MRTN-CT-2004-512385), the South-Eastern Norway Regional Health Authority and the Research Council of Norway are gratefully acknowledged. The authors thank Bjørg Simonsen for conducting the HLA-DQ2 peptide binding assay.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2011.01.057.

References and notes

1. Sollid, L. M.; Markussen, G.; Ek, J.; Gjerde, H.; Vartdal, F.; Thorsby, E. *J. Exp. Med.* **1989**, *169*, 345.
2. Todd, J. A.; Bell, J. I.; McDevitt, H. O. *Nature* **1987**, *329*, 599.
3. Thorsby, E. *Hum. Immunol.* **1997**, *53*, 1.
4. Lundin, K. E.; Scott, H.; Hansen, T.; Paulsen, G.; Halstensen, T. S.; Fausa, O.; Thorsby, E.; Sollid, L. M. *J. Exp. Med.* **1993**, *178*, 187.
5. Sollid, L. M. *Nat. Rev. Immunol.* **2002**, *2*, 647.
6. Sollid, L. M.; Thorsby, E. *Gastroenterology* **1993**, *105*, 910.
7. Xia, J.; Bergseng, E.; Fleckenstein, B.; Siegel, M.; Kim, C. Y.; Khosla, C.; Sollid, L. M. *Bioorg. Med. Chem.* **2007**, *15*, 6565.
8. Kapoerchan, V. V.; Wiesner, M.; Overhand, M.; van der Marel, G. A.; Koning, F.; Overkleeft, H. S. *Bioorg. Med. Chem.* **2008**, *16*, 2053.
9. Thorsby, E.; Lie, B. A. *Transpl. Immunol.* **2005**, *14*, 175.
10. Vartdal, F.; Johansen, B. H.; Friede, T.; Thorpe, C. J.; Stevanovic, S.; Eriksen, J. E.; Sletten, K.; Thorsby, E.; Rammensee, H. G.; Sollid, L. M. *Eur. J. Immunol.* **1996**, *26*, 2764.
11. van de Wal, Y.; Kooy, Y. M.; Drijfhout, J. W.; Amons, R.; Papadopoulos, G. K.; Koning, F. *Immunogenetics* **1997**, *46*, 484.
12. Arentz-Hansen, H.; Körner, R.; Molberg, Ø.; Quarsten, H.; Vader, W.; Kooy, Y. M.; Lundin, K. E.; Koning, F.; Roepstorff, P.; Sollid, L. M.; McAdam, S. N. *J. Exp. Med.* **2000**, *191*, 603.
13. Arentz-Hansen, H.; McAdam, S. N.; Molberg, Ø.; Fleckenstein, B.; Lundin, K. E.; Jørgensen, T. J.; Jung, G.; Roepstorff, P.; Sollid, L. M. *Gastroenterology* **2002**, *123*, 803.
14. Sjöström, H.; Lundin, K. E.; Molberg, Ø.; Körner, R.; McAdam, S. N.; Anthonsen, D.; Quarsten, H.; Noren, O.; Roepstorff, P.; Thorsby, E.; Sollid, L. M. *Scand. J. Immunol.* **1998**, *48*, 111.
15. Jüse, U.; van de, W. Y.; Koning, F.; Sollid, L. M.; Fleckenstein, B. *Hum. Immunol.* **2010**, *71*, 475.
16. Quarsten, H.; McAdam, S. N.; Jensen, T.; Arentz-Hansen, H.; Molberg, Ø.; Lundin, K. E.; Sollid, L. M. *J. Immunol.* **2001**, *167*, 4861.
17. Arntzen, M. O.; Osland, C. L.; Raa, C. R.; Kopperud, R.; Doskeland, S. O.; Lewis, A. E.; D'Santos, C. S. *Proteomics* **2009**, *9*, 1400.
18. Vader, W.; Stepniak, D.; Kooy, Y.; Mearin, L.; Thompson, A.; Van Rood, J. J.; Spaenij, L.; Koning, F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12390.