Crossreactivity to vinculin and microbes provides a molecular basis for HLA-based protection against rheumatoid arthritis

Jurgen van Heemst1,*, Diahann T. S. L. Jansen1,*, Savvas Polydorides2, Antonis K. Moustakas3, Marieke Bax1, Anouk L. Feitsma1, Diënée G. Bontrop-Elferink4, Martine Baarse1, Diane van der Woude1, Gert-Jan Wolbink5, Theo Rispens5, Frits Koning4, René R. P. de Vries4, George K. Papadopoulos6, Georgios Archontis2, Tom W. Huizinga1 & René E. Toes1

The HLA locus is the strongest risk factor for anti-citrullinated protein antibody (ACPA) rheumatoid arthritis (RA). Despite considerable efforts in the last 35 years, this association is poorly understood. Here we identify (citrullinated) vinculin, present in the joints of ACPA RA patients, as an autoantigen targeted by ACPA and CD4 T cells. These T cells recognize an epitope with the core sequence DERAA, which is also found in many microbes and in protective HLA-DRB1*13 molecules, presented by predisposing HLA-DQ molecules. Moreover, these T cells crossreact with vinculin-derived and microbial-derived DERAA epitopes. Intriguingly, DERAA-directed T cells are not detected in HLA-DRB1*13 donors, indicating that the DERAA epitope from HLA-DRB1*13 mediates (thymic) tolerance in these donors and explaining the protective effects associated with HLA-DRB1*13. Together our data indicate the involvement of pathogen-induced DERAA-directed T cells in the HLA-RA association and provide a molecular basis for the contribution of protective/predisposing HLA alleles.
Rheumatoid arthritis (RA) is a chronic autoimmune disease affecting synovial joints that can lead to severe disability. Pivotal pathophysiological insight has been obtained by the identification of anti-citrullinated protein antibodies (ACPA)\(^{1,2}\). These autoantibodies target proteins that have undergone a post-translational conversion of arginine to citrulline, catalysed by peptidylarginine deiminases (PAD enzymes)\(^3\). ACPA are highly specific for RA, enriched in the joints of patients and can crossreact between citrullinated antigens that are expressed in the inflamed joints\(^4–8\). It is clear now that RA represents two main syndromes, ACPA\(^+\) and ACPA\(^−\) disease, each with distinct genetic and environmental risk factors and disease outcome\(^9–12\). Characteristics of ACPA (for example, isotype usage and epitope spreading) indicate the involvement of CD4\(^+\) T cell help in shaping the ACPA response\(^13\).

The most important genetic risk factor for ACPA\(^+\) RA is the HLA class II locus and risk is confined to a region with genes encoding for the beta chain of HLA-DR and the alpha and beta chain of HLA-DQ that are in tight linkage disequilibrium (LD) and inherited in haplotypes\(^14,15\). An understanding of the HLA class II association and the relative contribution of the HLA-DR and HLA-DQ locus has been lacking for the last 35 years. Next to the association of the predisposing alleles to ACPA\(^+\) RA, other HLA molecules are associated with protection. These protective HLA alleles, mainly HLA-DRB1*13, carry the five amino-acid sequence DERAA at positions 70–74 of the beta chain and protect also in the presence of predisposing alleles\(^16,17\). Intriguingly, protection by these alleles is transferred from mother to child, supporting an active protective role of these alleles, possibly via microchimeric cells influencing thymic selection of CD4\(^+\) T cells, and indicating a dominant role of HLA-DRB1*13 in disease protection\(^18\).

HLA-derived peptides are a dominant peptide source presented by HLA class II molecules. We therefore proposed that the protective effect of HLA-DRB1*13 is explained by presentation of an HLA-DRB1*13-derived peptide in the context of other (predisposing) HLA class II molecules\(^19–21\).

It was previously shown that the degradation of HLA-DRB1*13 can result in the presentation of a peptide with the core sequence DERAA by other HLA class II molecules to CD4\(^+\) T cells\(^22\). This could allow for the negative selection of such 'DERAA-directed' CD4\(^+\) thymocytes. Interestingly, the DERAA sequence is also found in many microbes and in the self-protein vinculin. Vinculin is expressed in the synovium and was recently shown to be citrullinated in the synovial fluid of an RA patient\(^2,3\). Likewise, T cells directed to vinculin are found under certain infectious conditions, indicating that T cell tolerance to vinculin is not absolute\(^24,25\). Molecular mimicry of self-proteins with pathogenic proteins was proposed as an important mechanism to break T cell tolerance\(^26,27\). Therefore, we postulate that on priming of DERAA-directed T cells by microbes expressing DERAA-containing proteins, T cells crossreactive to vinculin would be able to provide help to B cells reactive to citrullinated vinculin. This would ultimately result in the production of ACPA. In HLA-DRB1*13-positive donors, these T cells are conceivably deleted, leading to protection against ACPA\(^+\) disease.

Here we show that citrullinated vinculin is novel autoantigen for ACPA antibodies. In addition, we demonstrate the presence of a T cell population in HLA-DRB1*13-negative donors that specifically recognize a DERAA-containing vinculin epitope and that crossreact with DERAA sequences derived from pathogens.

**Results**

Citrullinated vinculin is a novel target for ACPA. We speculate that the protective effects of HLA-DRB1*13 on the development of ACPA\(^+\) RA are related to a T cell response reactive to a sequence that is commonly present in the HLA-DRB1*13 molecule, microorganisms and self-proteins that are targeted by ACPA. Indeed, a 5-amino-acid long HLA-DRB1*13-derived sequence (DERAA) is present in many microorganisms and a few self-proteins. Of these self-proteins, the citrullinated protein vinculin attracted our attention. Therefore, we first analysed whether citrullinated vinculin is recognized by ACPA. To this end, we citrullinated vinculin in vitro with PAD enzymes and tested both native and citrullinated vinculin for recognition with serum of an ACPA\(^+\) RA patient. We observed citrulline-specific recognition of vinculin by serum of an ACPA\(^+\) RA patient, but not with an ACPA\(^−\) patient (Fig. 1a). We further confirmed that citrullinated vinculin is a target of ACPA using an ACPA monoclonal antibody (Fig. 1b)\(^28\).

Recently, several reports showed the presence of citrullinated vinculin in the synovial fluid of ACPA\(^+\) RA patients. In addition, three sites of in vivo citrullination on this protein were identified: Arg285, Arg622 and Arg823 (refs 4,23). When we studied antibody responses to these citrullination sites, we could show recognition by 34% of tested sera from ACPA\(^+\) RA patients versus 5% of sera from ACPA\(^−\) patients (Fig. 1c). Sera of ACPA\(^+\) RA patients are highly (cross-)reactive towards multiple citrullinated antigens. Indeed, when we quantified responses to citrullinated alpha-enolase, fibrinogen, vimentin and myelin basic protein in patient sera that react to citrullinated vinculin peptides, we could readily demonstrate additional reactivities indicating...
that ACPA are not exclusively directed against citrullinated vinculin as expected (Fig. 1d).

Together these data show that citrullinated vinculin is a self-protein recognized by RA autoantibodies in a citrulline-dependent fashion.

**Vinculin is recognized by T cells from HLA-DRB1*13 negative donors.** We next wished to determine whether the DERAA sequence from vinculin is recognized by human CD4+ T cells. To this end, peripheral blood mononuclear cells (PBMCs) were stimulated with a 15-mer vinculin-peptide (VCL622–636, REEVFDERAANH) (VCL-DERAANH) (VCL-DERAA) for 24 h. We observed a clear reactivity towards this epitope as determined in an interferon-γ (IFN-γ)-ELISPOT assay (Supplementary Fig. 1). HLA-DRB1*13, especially HLA-DRB1*13:01, is strongly associated with protection against ACPA+ RA. We hypothesize that HLA-DRB1*13 protects against ACPA+ disease by affecting the generation of DERAA-directed T cells. Therefore, we stratified donors for HLA-DRB1*13:01 status. We observed a striking difference in IFN-γ-producing cells depending on HLA-DRB1*13:01-status. The lack of reactivity in HLA-DRB1*13:01-positive donors was not due to a hampered ability of such donors to respond to T cell antigens as we observed strong response to microbial antigens in both HLA-DRB1*13:01 carriers and non-carriers when stimulated with recall antigens (Fig. 2a). We further confirmed this finding in PBMCs stimulated for 4 days (Supplementary Fig. 2). The differential ability of HLA-DRB1*13:01-positive donors to respond to VCL-DERAA is most likely not due to a general deficiency to present the VCL-DERAA epitope as these donors were heterozygous and thus expressed other HLA molecules that could potentially present VCL-DERAA.

Nonetheless, to further confirm that HLA-DRB1*13:01 affects the ability to generate VCL-DERAA-directed T cell responses, we repeated the experiments in a set-up stratified for HLA using PBMCs from HLA-DRB1*04-, HLA-DRB1*13:01- and HLA-DRB1*04/*13:01 heterozygous donors. A significant reduction in IFN-γ-producing cells was observed in both HLA-DRB1*13:01 carriers and DRB1*04/*13:01 heterozygous donors as compared with HLA-DRB1*04 carriers. Again, no difference was observed for recall antigens (Fig. 2b). These data indicate that the lack of detecting VCL-DERAA-directed T cells is not explained by the inability of HLA-DRB1*13 molecules to present VCL-DERAA, but rather the result of a dominant effect associated with the presence of HLA-DRB1*13.

Thus, vinculin is an autoantigen recognized by circulating VCL-DERAA-directed CD4+ T cells. In HLA-DRB1*13:01 carriers these T cells are absent. This effect, like the protective effects of HLA-DRB1*13:01 on the development of ACPA+ RA, was present in a dominant fashion, consistent with the notion that HLA-DRB1*13 affects the generation of DERAA-directed T cells, possibly during thymic selection.

**Presentation of Vinculin-DERRA by RA-predisposing HLA-DQ molecules.** HLA-DRB1 alleles predisposing to ACPA+ RA in the Caucasian population are collectively called HLA shared epitope (SE) alleles, The most common HLA-SE alleles are HLA-DRB1*04:01 and HLA-DRB1*01:01. Using HLA-DR-binding assays, we examined the ability of VCL-DERRA to bind to these HLA molecules, but we could not detect any binding of VCL-DERRA to these alleles (Table 1, Supplementary Fig. 3a,b). The HLA class II region is known for its strong LD. Genes that encode for the HLA-DR-beta chains are inherited in haplotypes with genes that encode for the alpha and beta chain of HLA-DQ. The predisposing HLA-SE alleles are in tight LD with HLA-DQ5 (DQA1*01;DQB1*05:01), HLA-DQ7.3 (DQA1*03:02; DQB1*03:01) and HLA-DQ8 (DQA1*03:01;DQB1*03:02) (Fig. 3a,b). In Fig. 3b, the distribution of HLA-DQ molecules in SE+ healthy donors is depicted, showing that these donors are HLA-DQ5-, HLA-DQ7.3- or HLA-DQ8 positive. We therefore

![Figure 2](https://example.com/figure2.png) **Figure 2 | Citrullinated vinculin is a novel autoantigen targeted by CD4+ T cells in HLA-DRB1*13-negative donors.** (a) IFN-γ ELISPOT of PBMCs from healthy individuals stimulated for 24 h with VCL-DERRA peptide REEVFDERAANFENH (n = 21) or with a mix of recall antigens (tetanus, tuberculin, candida, n = 18). Donors were stratified based on HLA-DRB1*13:01 status. (b) IFN-γ ELISPOT of PBMCs of HLA-DRB1*13:01 carriers (n = 7), HLA-DRB1*04 carriers (n = 12) and HLA-DRB1*13:01/DRB1*04 donors (n = 9), stimulated for 24 h with VCL-DERRA peptide or with a mix of recall antigens. Each dot represents a unique donor. Two-sided statistical analyses of ELISPOT data were performed using a Mann-Whitney U-test with * indicating P < 0.05. Neg, negative; NS, not significant; Pos, positive.

| Association | Subtype | HLA allele | IC50 (μM) |
|-------------|---------|------------|-----------|
| Susceptibility haplotypes | HLA-DR | DRB1*01:01 | >300 |
| | HLA-DR | DRB1*04:01 | >300 |
| | HLA-DQ | DQ5 | 39 |
| | HLA-DQ | DQ7.3 | 11 |
| | HLA-DQ | DQ8 | 14 |
| Non-susceptibility haplotypes | HLA-DR | DRB1*13:01 | >300 |
| | HLA-DR | DRB1*13:02 | >300 |
| | HLA-DQ | DQ6:2 | >300 |
| | HLA-DQ | DQ6:3 | >300 |
| | HLA-DQ | DQ6:4 | >300 |
| | HLA-DQ | DQ7:5 | >300 |

Affinity of VCL-DERRA for associated HLA class II molecules. Affinity was determined by a competitive binding assay using high-affinity biotinylated peptides. Results are depicted as IC50 value, the concentration of test peptide (μM) where 50% of biotinylated peptide is bound. Experiments were performed at least three times and the average IC50 value between the different experiments is shown.
next studied the binding of VCL-DERAA to these HLA-DQ alleles and observed binding of the VCL-DERAA peptide to HLA-DQ5, DQ7.3 and DQ8 (Table 1, Supplementary Fig. 3c–e). These data indicate that VCL-DERAA can be presented by HLA-DQ molecules in LD with predisposing HLA-SE alleles.

To obtain an indication whether the VCL-DERAA can be presented by more HLA molecules, we also studied the ability of additional HLA-DR and HLA-DQ molecules encoded by HLA haplotypes that protect or have no influence on the development of ACPA+ RA. We could not detect binding of VCL-DERAA to these HLA class II alleles (Table 1, Supplementary Fig. 3f–k). Likewise, using an IFN-gamma ELISPOT, VCL-DERAA-directed T cells were absent in donors negative for predisposing HLA-DQ alleles (n = 31).

**Figure 3 | Presentation to vinculin-DERAA-directed T cells is restricted to RA-predisposing HLA-DQ molecules. (a) Schematic representation of the LD between the HLA-DRB1, HLA-DQA1 and HLA-DQB1 gene in studied haplotypes associated or not associated with susceptibility to ACPA+ RA. (b) Distribution of HLA-DQ molecules in ACPA+ RA patients positive for HLA-DR SE alleles (n = 31).**

Together HLA class II presentation and T cell recognition of VCL-DERAA was restricted to HLA-DQ molecules that are associated with RA susceptibility.

**Identification of microbe-specific T cells targeting DERAA epitopes.** We next investigated the presence of the DERAA sequence in microorganisms. A Blast search showed that DERAA is found in 66% of bacteria and 4% of viruses (Fig. 4a). This large number represents a major challenge to identify potential ‘crossreactive microorganisms’. To select for relevant microorganisms, we restricted the search to those microorganisms that can cause disease or symptoms in humans and are present in the western world. This approach left us with 219 candidate sequences. We then synthesized eight DERAA-containing peptides from common recall microbes including several bacteria (*P. acnes, S enteritidis, B. pertussis, S. aureus*) and viruses (measles virus, influenza A and human herpesvirus 7). Interestingly, all of these peptides were presented by HLA-DQ7.3 and DQ8 molecules showing that these HLA molecules can efficiently accommodate both the VCL-DERAA epitope as well as microbe-derived DERAA epitopes (Fig. 4b).

To determine whether these microbial-derived epitopes are recognized by human T cells, we generated T cell lines by stimulating PBMCs from three healthy HLA-DQ8-positive donors for 7 days with a pool of the eight pathogen-derived peptides (PathMix). Next we determined the presence of PathMix-specific CD4+ INF-γ-producing T cells. As shown in Fig. 4c, such T cells were readily detectable. Using limiting dilutions, we isolated T cell clone D2C18, further confirming the presence of ‘microbial-DERAA-directed’ T cells (Supplementary Fig. 9). Next we also analysed the presence of ‘microbial-DERAA-directed’ T cells by ELISPOT directly *ex vivo*, allowing the analyses of more donors at higher throughput. Interestingly, PBMCs from HLA-DRB1*13:01 carriers displayed a significantly reduced reaction against these peptides, indicating that the presence of HLA-DRB1*13 also affects the formation of T cell responses against DERAA epitopes from microbes (Fig. 4d).

Together these data indicate that HLA-DQ alleles associated with ACPA+ RA efficiently present both VCL-DERAA and microbe-derived DERAA epitopes and that microbe-specific T cells directed to DERAA epitopes are readily detected. The presence of HLA-DRB1*13:01 affects the formation of these T cell responses.

**Predicting crossreactive microbes by modelling vinculin-DERAA presentation.** The data presented above demonstrate the presence of microbial- and VCL-DERAA-directed T cells, but do not show if a single T cell receptor (TCR) can react with both epitopes. To facilitate the search for possible ‘crossreactive’ DERAA epitopes out of all microbe-derived DERAA sequences, we first determined the binding register of VCL-DERAA using HLA class II-binding assays with amino (N)- and carboxyl (C)-terminally truncated and amino-acid-substituted VCL-DERAA peptides as detailed in the Supplementary Note and in Supplementary Figs 10–14. Together these experiments support VFDERAANF (anchors in bold) as the core binding register for both HLA-DQ7.3 and HLA-DQ8. All-atom molecular dynamics (MD) simulations showed that residues Val625 (P1), Glu628 (P4) and Phe633 (P9) make numerous intermolecular polar and nonpolar interactions in the respective pockets. In Fig. 5a–f, we highlight key intermolecular interactions between the VCL-DERAA epitope and the HLA-DQ8 molecule. A detailed discussion of the MD results and a quantitative assessment of the intermolecular interactions can be found in the supplementary data (Supplementary Figs 15 and 16, Supplementary Table 1 and
Supplementary Note). The MD simulations also indicated that the long protruding side chains of Glu623 (P-2) Glu624 (P-1), Phe626 (P2), Asp627 (P3), Arg629 (P5), Asn632 (P8), Glu634 (P10) and Asn635 (P11) are exposed and could potentially interact with crossreactive TCRs (Fig. 5g). Subsequently we also further confirmed the obtained model using a second type of molecular modelling: energy minimization, which further confirmed possible TCR-contact residues (Supplementary Figs 17 and 18).

To functionally confirm whether a TCR would indeed interact with (some of the) potential TCR-contact residues within the VCL-DERAA peptide, we next determined how the epitope is recognized by the VCL-DERAA-directed T cell clone JPT57. Phe626 (P2) interacts with the JPT57–TCR as its removal results in a large decrease in recognition without affecting HLA-DQ8 binding. (Fig. 5h, Supplementary Fig. 12). C-terminal truncations resulted in a large decrease in T cell recognition after the removal of Glu634 (P10), showing that this residue is also important for JPT57–TCR interaction (Fig. 5i). Thus, the data obtained using the VCL-DERAA-directed T cell clone as a functional read-out, are in line with the HLA-binding and molecular modelling studies indicating the sequence VFDERRAA (anchors in bold) as the minimal epitope required for activation of JPT57.

Next we performed alanine substitutions within the minimal epitope to remove critical TCR-interacting residues. Substituting Asn632 (P8) and Phe633 (P9) dramatically impacted T cell recognition, without affecting the binding affinity of VCL-DERAA (Fig. 5j,k).

Together these data indicate VFDERRAA (anchors in bold) as the most likely core binding register for HLA-DQ7.3 and HLA-DQ8 and Asn632 and Phe633 as important residues for JPT57–TCR recognition.

Figure 4 | Identification of pathogen-specific T cells targeting DERAA epitopes in HLA-DRB1*13 negative donors. (a) Frequency of bacteria or viruses positive for a DERAA-containing protein as identified using the NCBI BLAST. List of all bacteria or viruses in blast database was generated and for three random sets of 50 bacteria/viruses a blast search for DERAA was performed. Error bars indicate the s.d. between the three analysed sets. (b) Competitive binding of pathogen-derived DERAA epitopes to HLA-DQ7.3 or HLA-DQ8. Binding experiments were performed at least three times and plots show pooled experiments, the error bars show the variation between the different experiments. (c) Flow cytometry staining of PathMix-directed T cell lines restimulated with PathMix pulsed or unpulsed antigen-presenting cells. Plots are gated on CD14neg/CD4pos/CD25pos T cells. (d) ELISPOT of PBMCs of healthy individuals of HLA-DRB1*13:01 donors (n = 5) and HLA-DQ7.3/DQ8 donors (n = 8) stimulated for 24 h with DERAA-containing peptides and recall antigens. Error bars indicate the variation between different donors. Error bars are s.e.m.
TCR crossreactivity between vinculin and bacterial antigens. Identifying microbes that crossreact with vinculin was challenging due to the large number of potential candidates. Therefore, we used the data presented above to determine if a single TCR can crossreact both to DERAA sequences from microbes and the self-protein vinculin.

As we identified the Asn632 and Phe633 at P8 and P9, respectively, as important for JPT57–TCR interactions, we used VCL-DERAA to further investigate the interaction of the TCR with vinculin.

**Figure 5** | Molecular modelling of HLA:VCL-DERAA complexes identifies critical TCR-contact residues. (a–c) Non-polar contacts between the VCL-DERAA peptide and HLA-DQ8 in pockets P1 (a), P4 (b) and P9 (c). Protein residues are depicted in a surface representation, in distinct colours. (d–f) Hydrogen bonds between the VCL-DERAA peptide and HLA-DQ8 in pockets P1 (d), P4 (e) and P9 (f). Peptide and protein groups are shown, respectively, in thick and thin licorice representation and water molecules as small spheres. Hydrogens are omitted, with the exception of the extra hydrogen in protonated αGlu31. (g) TCR view of the DQ8:VCL-DERAA complex. The peptide main chain is shown in licorice; domains α and β are shown in a surface representation, coloured, respectively, in blue and red. Side chains P-2E, P-1E, P2F, P3D, P5R, P8N, P10E and P11N are facing towards the TCR-binding site. The P2F and αPhe58 side chains and the non-polar moiety of P5R form a non-polar cluster, which faces towards the TCR-binding site; the average P2F ring–αPhe58 ring distance is ~8 Å. (h–j) IFN-γ ELISA on supernatant of T cell clone JPT57 stimulated with unpulsed HLA-DQ8-positive antigen-presenting cells and N-truncated (h), C-truncated (i) or alanine-substituted (j) VCL-DERAA peptides. (k) Competitive binding of the VCL-DERAA peptide, and alanine-substituted VCL-DERAA peptides to HLA-DQ8. Binding experiments were performed at least three times and plots show pooled experiments, the error bars show the variation between the different experiments. Experiments involving JPT57 were performed three times and show a representative experiment. Two-sided statistical analysis was performed using a Student’s t-test with ** indicating P<0.001. Error bars are s.e.m.
the molecular modelling to predict which viral and bacterial epitopes are likely to bind to HLA-DQ8 and harbor P8Asn or P9Phe and are thus likely to crossreact with JPT57. From the 15 peptides identified (Supplementary Table 2), three were able to activate JPT57 in a T cell stimulation assay at concentrations similar to those used for the vinculin peptide. These epitopes were derived from Campylobacter coli, Lactobacillus curvatus and Lactobacillus sakei and crossreacted with JPT57 in an HLA-DQ-dependent manner (Fig. 6a). Peptide-binding studies revealed that these three peptides bind with an intermediate binding affinity to HLA-DQ7 and HLA-DQ8 (Fig. 6b,c). Molecular models of these bacterial peptides in HLA-DQ8 illustrate the similarities with VCL-DERAA (Fig. 6d–g, Supplementary Figs 19–21).

Thus, these data show that the RA-susceptibility alleles HLA-DQ7 and HLA-DQ8 can present both VCL-DERAA and related microbe-derived epitopes to T cells and that such T cells can be crossreactive to vinculin and bacterial epitopes, thereby providing an explanation for the presence of activated self-reactive CD4+ T cells directed to vinculin in peripheral blood.

### Discussion

The strong connection between the HLA locus and RA has been known for more than 35 years. The complex HLA class II associations and the diverse ACPA responses in RA patients suggest the presence of multiple aetiological pathways. To unravel these pathways, identifying the relevant autoantigens is crucial. We here present evidence favouring the involvement of vinculin in the emergence of ACPA+ disease. This cytoskeletal protein was recently found to be citrullinated in vivo in the synovial fluid. We now show that it is recognized by ACPA as well, thereby adding it to a still selective list of targets. Moreover, we identified an epitope from vinculin recognized by CD4+ T cells restricted to HLA-DQ7 and HLA-DQ8 molecules predisposing to ACPA+ RA. The core amino-acid sequence (DERAA) is also present in many pathogens and in HLA-DRB1*13, a molecule encoded by an HLA locus associated with protection against ACPA+ RA. We have also shown that a single TCR can recognize both a vinculin-DERAA epitope as well as DERAA epitopes from microbes, indicating the crossreactive nature of ‘DERAA’-directed T cell responses. More importantly, such T cell responses appear absent from donors carrying HLA-DRB1*13 as DERAA-directed T cell responses to either pathogen- or vinculin-derived DERAA epitopes were lacking in these subjects. Even donors that harboured HLA-DRB1*13 next to predisposing HLA alleles were unable to respond to vinculin-DERAA.

Together these data indicate a novel pathway that explains several of the protective and predisposing HLA-effects associated with ACPA+ RA (Fig. 7). In short, recognition of citrullinated vinculin by B cells will lead to the presentation of the ‘VCL-DERAA’ epitope in the context of HLA class II molecules. The HLA-DQ molecules genetically linked to the predisposing HLA-SE-molecules are particularly good in presenting DERAA-containing peptides. The DERAA-directed T cells primed against such peptides crossreact with the VCL-DERAA peptide and provide help to the B cells, ultimately leading to a strong ACPA response. Subjects born with HLA-DRB1*13, will present the HLA-DRB1*13-derived DERAA-peptide in the thymus, leading to tolerization of the
HLA-DRB1*13 negative

Thymus

**DERAA-directed T cell escape**

Negative selection

HLA-DRB1*13 positive

Thymus

**DERAA-directed T cell**

Infection

**T cell activation**

Vinculin-derived DERAA peptide

HLA-DR13-derived DERAA peptide

Microbial protein-derived DERAA peptide

**Molecular mimicry**

**B cell activation**

ACPA

**Synovial joint**

**Inflammation**

**Synovium**

**Figure 7 | Schematic representation of the proposed role of DERAA-directed CD4+ T cells in ACPA+ RA.** DERAA-directed T cells are restricted to RA-predisposing HLA-DQ5, -DQ7.3 and -DQ8 molecules. In carriers of these HLA-DQ molecules, DERAA-directed T cells can become activated on contact with microbes. These activated T cells can subsequently crossreact with a DERAA epitope derived from vinculin resulting in the activation of citrulline-directed B cells and the production of ACPA directed to citrullinated vinculin or vinculin-linked proteins. Citrullinated vinculin is present in the synovial compartment and is a target of ACPA. Binding of ACPA to its target can induce antibody-mediated effector mechanisms thereby contributing to synovial inflammation. In HLA-DRB1*13-positive individuals, a HLA-DRB1*13-derived DERAA epitope is presented by predisposing HLA-DQ molecules to CD4+ T cells resulting in their negative selection, thereby protecting against the development of ACPA+ RA.

DERAA-reactive T cell response and hence the inability to provide help to ACPA-producing B cells via this pathway and thereby the emergence of ACPA+ RA (Fig. 7).

The variation of HLA-DR and HLA-DQ molecules in the human population is enormous. We have shown that predisposing HLA-DQ molecules are particularly good at presenting the VCL-DERAA epitope. Interestingly, the absence of VCL-DERAA affinity for a wide variety of other tested HLA-DR or HLA-DQ molecules could indicate a selective presentation by these HLA risk molecules. Next to a role of predisposing haplotypes, we also focussed on the protective effect of HLA-DRB1*13 alleles. However, the DERAA sequence can also be found in other HLA-DRB1 alleles (*04:02, *11:02, *11:03), which are rare in Caucasian populations. These alleles have previously been implicated in protection from ACPA+ RA, but their allele frequency hampers functional studies. Interestingly, it has been previously reported that the processing of these alleles results in the generation of a similar HLA-DERAA epitope suggesting that these alleles could all protect via the pathway that we have described.

The place in time at which HLA-DR13 mediates protection from the development of ACPA and/or ACPA+ disease, or its relation to epitope spreading of the ACPA response is currently not known and would be relevant to determine in future studies. Recent evidence showed that the ACPA response matures before disease onset and that the HLA system could be involved in this RA connection, including both predisposing and protective HLA effect, and warrants further studies addressing the possibility to target DERAA-directed T cells in the prevention of ACPA+ RA.

**Methods**

**Cells and sera.** HLA-typed buffy coats from healthy volunteers were obtained from the blood bank (Sanquin, The Netherlands). PBMCs and sera from RA patients were derived from patients participating in the Leiden Early Arthritis Clinic cohort. All RA patients fulfilled the American College of Rheumatology criteria for the classification of RA. Patient samples were compared with 80 control samples from healthy individuals also living in the Leiden area. PBMCs were isolated using a standard Ficoll procedure. The protocols were approved by the Leiden University Medical Center ethics committee and informed consent was obtained.

**Peptides.** Peptides were synthesized according to standard Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry using a SyroII peptide synthesizer (MultiSynTech, Witten, Germany). The integrity of the peptides was checked using reverse-phase high-performance liquid chromatography and mass spectrometry. For vinculin ELISPOT experiments, HLA class II-binding studies and T cell activation studies 15-mer vinculin epitope VCL622–636 (VCL-DERAA;
REEVE DERAANFENH) was used. To determine the peptide-binding register, the VCL622–636 peptide was modified. Truncated on the N- or C-terminus or alanine-substituted at different positions. For ELISPOT assays against common pathogens, pathogenic peptides were used derived from measles virus (SSRASADERAFLPS), influenza A (VFEFSDEARAAPQV), human herpesvirus 7 (LAARADERAAPPDVG), Bordetella pertussis (SPNLTDER AAQAGVT), Staphylococcus aureus (QMDNDAERALTAM), Haemophilus influenzae (RFHGDDERAAKV) and Salmonella enteritidis (PMLMDR AAQKYEN) and Propionibacterium acet (EVEPTDERRALSHV). For the generation of these T cells, pathogenic peptides were synthesized.

**Cell lines.** The following Epstein–Barr virus-transformed lymphoblastoid B-cell lines were used: BSM (DRB1*04:01;DQA1*03:01;DQB1*03:02, IHW902), BOLETH (DRB1*04:01;DQA1*03:01;DQB1*03:02, IHW903), BM21 (DQA1*05:01, IHW901), ISM (DQA1*03:01, IHW901), LAP (DQA1*03:01;DQB1*03:06, IHW9291), BC34 (DRB1*13:02;DQA1*01:02;DQB1*06:04), WT8 (DQA1*01:02;DQB1*06:02, IHW9017), KAS116 (DRB1*01:01;DQA1*01:02;DQB1*05:01, IHW9003) and 72:1.2 (DR/DQ-negative)38. Cells were maintained in IMDM (Lonza) supplemented with 10% heat-inactivated fetal calf serum. JSM and BC34 cells are derived from healthy HLA-typed donors that were Epstein–Barr virus transformed.

**HLA class II competitive peptide-binding assay.** Peptide-binding assays were performed, as described previously39. In short, cell lysates from HLA class II homozygous B-lymphoblastoid cell lines were incubated on SPV3- (anti-HLA-DQ) or B8.11.2- (anti-HLA-DR) coated (10 μg/ml) Fluoronec 96-well plates at 4°C overnight. Titration ranges of the tested peptides (0 to 300 μM) were mixed with a 3:1:0.6 μM of non-concentration indicator peptide and added to the wells. Bound indicator peptide was detected using europium–strepavidin (Perkin Elmer, Boston, MA) and measured in a time-resolved fluorometer (PerkinElmer, Wallac Victor2). IC50 values were calculated based on the observed binding of the test peptide against the fixed concentration indicator peptide. The IC50 value describes the concentration of test peptide required for a loss of 50% of the indicator peptide signal.

**Enzyme-linked immunosorbent spot assay.** PBMCs were incubated in 96-well Multiscreen HA plates (Millipore) precoated with 5 μg/ml anti-IFN-γ capture antibody (clone 1-D1K; Mabtech) at 5 ×106 cells per well and stimulated in X-VIVO medium (Lonza) with 5 μg/ml peptide or with a mix of recall anti-CD3/CD28 (Clone OKT3, Mabtech) antibody and peptide added to the wells. Bound indicator peptide was detected using europium–strepavidin (Perkin Elmer, Boston, MA) and measured in a time-resolved fluorometer (PerkinElmer, Wallac Victor2). IC50 values were calculated based on the observed binding of the test peptide against the fixed concentration indicator peptide. The IC50 value describes the concentration of test peptide required for a loss of 50% of the indicator peptide signal.

**Flow cytometry.** Polyclonal CD4+ T cell lines were generated by stimulating 3 ×106 PBMCs per well in a 24-well plate with 5 μg/ml peptide for 7 days in IMDM supplemented with 5% human serum (Sanquin). After 7 days, 1.5 × 105 autologous PBMCs were plated in six-well plates for 2 h. Non-adherent cells were removed and adherent cells were pulsed with 5 μg/ml peptide and used as feeders for 106 T cells. After 1 h, 10 μg/ml brefeldin A was added. Cells were incubated overnight and used for intracellular cytokine staining. The cells were incubated with fluorochrome-conjugated antibodies recognizing CD4 (Clone RPA-T4; BD biosciences), CD45 (Clone 1D3; eBioscience) and CD25 (Clone M-A251; BD biosciences), after which they were permeabilized using CytoFix/Cytoperm Kit (BD Biosciences). After washing, cells were incubated with PE-labelled anti-IFN-γ and matching isotype control. Cells were taken up in 1% paraformaldehyde until flow cytometric acquisition. Flow cytometry was performed on FACS Calibur (BD biosciences) or LSR II (BD biosciences). Analysis was performed using FACS Diva (BD biosciences) and Flowjo software.

**T cell cloning.** JPT57 was generated from an HLA-DRB1*04:01;DQA1*03:01;DQB1*03:02, IHW902 and JPT57 was generated from an HLA-DRB1*04:05/01:01;DQ8/DQ5 generating mix of recall anti-CD3/anti-CD28. JPT57 was generated from an HLA-DRB1*04:05/01:01;DQ8/DQ5 generating mix of recall anti-CD3/anti-CD28. JPT57 was generated from an HLA-DRB1*04:05/01:01;DQ8/DQ5 generating mix of recall anti-CD3/anti-CD28.
peptides are depicted in van der Waals surface representation, with colour and
depiction conventions identical to those for the other figures. Several visible
residues of the HLA-DQ molecule in contact with the antigenic peptide and
potential contact with a cognate TCR in canonical orientation are shown in stick
form with a transparent surface (atomic colour code: oxygen, red; nitrogen, blue;
hydrogen, white; carbon, orange; sulfur, yellow). The antigenic peptide in the
groove is shown in space filling form with the same atomic colour code as in Fig.
6d-g. Coordinates of all complexes shown in the various figures have been deposited in the
Fibsgare repository under accession codes 1294716.

MD simulations. The simulated system consisted of the entire DQ8 molecule and
the 13-residue peptide with the vinculin sequence Glu-Glu-Val-Phe-Asp-Glu-Arg-
Ala-Ala-Asn-Phe-Glu-Asn. Titratable residues were assigned their most
common ionization state at physiological pH, with the exception of DQ8 residue sGlu31,
which was protonated. In the crystallographic structure of the DQ8-insulin com-
plex, residues sGlu31 and sGlu86 are in direct contact and their geometry and
interactions with nearby residues and a crystallographic water suggest that at
least one of them is protonated. A similar conclusion was reached for a pair of Glu and
Asp residues in the pocket P6 of DR and I-E molecules45. We used the empirical
model Propka42 and a constant-pH Monte Carlo approach implemented in the
program PROTEUS46 to compute the pK of titratable groups in the
crystallographic structure of HLA-DQ8, both in the absence and the presence of
the peptides insulin and VCL-DERRA. Both methods agreed that the residues
sGlu31 and sGlu86, near pocket P1, are strongly correlated; the predicted pKa values
suggested that one of them should be protonated, most probably sGlu31.
The initial coordinates of the protein heavy atoms were taken from the
crystallographic structure of the DQ8-insulin complex (PDB accession code 1JK8)42. The
peptide main chain heavy atoms were placed at the corresponding coordinates
of the insulin main chain. Hydrogens were positioned by the HBUILD
algorithm of the CHARMM programme47. The peptide side chain initial
conformations were optimized with the program PROTEUS45. An additional
control simulation studied the DQ8insulin complex; for this system, the initial
coordinates of the protein and peptide were taken from the corresponding
crystallographic structure (PDB accession code 1JK8)42.
The initial set-up of the simulation system was performed with the CHARMM-
GUI interface48. A total of 70 crystallographic waters of the DQ8:insulin complex
were retained. For the vinculin complex, crystallographic waters were minimized
for 100 steps before the simulation, with the protein and peptide atoms kept fixed.
The ligand complexes were immersed in a periodically replicated water box with the
shape of a truncated octahedron. Overlapping water molecules were
omitted and 17 potassium ions were added (14 ions in the insulin complex), to
neutralize the total charge. The final complex had 120,324 atoms (6,167 protein-
ligand atoms); the insulin complex had 120,504 atoms (6,172 protein-ligand
atoms).

The simulations employed the molecular mechanics program CHARMM c37b2
(ref. 47). Protein atomic parameters were taken from the CHARMM36 all-atom
core force field with a CMAP backbone phi/psi energy correction95,50. Water
parameters corresponded to the modified TIP3P water model51,52. Electrostatic
interactions were calculated without truncation by the particle-mesh Ewald
algorithm53,54. van der Waals interactions were calculated without truncation by the
particle-mesh Ewald

\[
\Delta G_{opt}^{\text{opt}} = \sum_{i<j} \left( \sigma_{ij} \right) + \sum_{i<j} \beta \left( V_{ij}^{\text{LJ}} + V_{ij}^{\text{MB}} \right) + \sum_{i<j} \frac{q_i q_j}{R_{ij}} + \sum_{i<j} \frac{q_i q_j}{R_{ij}} + \sum_{i<j} \frac{q_i q_j}{R_{ij}} \]

(1)

The first and second group of terms on the right-hand side of equation (1) describe,
respectively, polar and non-polar interactions between R and R; ‘Coul’ denotes Coulombic interaction, ‘GB’ denotes generalized born interaction, \(\Delta G_{opt}^{\text{opt}}\) is the change
in the solvent accessible surface area of atom on binding and \(\sigma\) is a surface tension
coefficient. The residue-pair interaction energies of equation (1) include solvent-
mediated effects via the above GB and surface area terms. In the calculations
of Supplementary Fig. 15, R and R’ are distinct ligand and protein residues.

We employed the GB/WS Generalized Born model55,56 as implemented in the
CHARMM programme. The coefficient \(\sigma\) was set to 0.005 kcal mol

-1 Å

2 for consistency with the GBSW parameterization. To compute the GB contributions,
we removed all waters and ions from the simulation system and set the charge to
zero for protein and peptide atoms other than those belonging to residues R and R’,
respectively. The last term contains the difference in solvent accessible surface areas
of groups R and R’ in the complex and unbound states. Coordinates of all complexes
shown in the various figures have been deposited in the Figsahare repository under accession code 1294716.

References
1. Schellekens, G. A., de Jong, B. A., van Hoogen, F. H., van de Putte, L. B. &
van Veenoo, W. J. Citrulline is an essential constituent of antigenic
determinants recognized by rheumatoid arthritis-specific autoantibodies.
J. Clin. Invest. 101, 273–281 (1998).
2. Girbal-Neuhauser, E. et al. The epitopes targeted by the rheumatoid arthritis-
associated antiflaggrin autoantibodies are posttranslationally generated on
several sites of (pro)flaggrin by deimination of arginine residues. J. Immunol.
162, 585–594 (1999).
3. Azzaliotti, C., Pratesi, T., Tommasi, C. & Migliorini, P. Pepsidylarginine
deiminase 4 and citrullination in health and disease. Autoimmun. Rev. 9,
158–160 (2010).
4. van Beers, J. J. et al. The rheumatoid arthritis synovial fluid citrullinome reveals
novel citrullinated epitopes in apolipoprotein E, myeloid nuclear differentiation
antigen, and beta-actin. Arthritis Rheum. 65, 69–80 (2013).
5. Vossenaaar, E. R. et al. The presence of citrullinated proteins is not specific for
rheumatoid synovial tissue. Arthritis Rheum. 50, 3485–3494 (2004).
6. Ioan-Facsinay, A. et al. Anti-cyclic citrullinated peptide antibodies are a
collection of anti-citrullinated protein antibodies and contain overlapping and
non-overlapping reactions. Ann. Rheum. Dis. 70, 188–193 (2011).
7. Amara, K. et al. Monoclonal IgG antibodies generated from joint-derived B
cells of RA patients have a strong bias toward citrullinated autoantigen
recognition. J. Exp. Med. 210, 445–455 (2013).
8. Snir, O. et al. Antibodies to several citrullinated antigens are enriched in the
joints of rheumatoid arthritis patients. Arthritis Rheum. 62, 44–52 (2010).
9. van Gaalen, F. A. et al. Autoantibodies to cyclic citrullinated peptides predict
progression to rheumatoid arthritis in patients with undifferentiated arthritis: a
prospective cohort study. Arthritis Rheum. 50, 709–715 (2004).
10. Rantapaa-Dahlqvist, S. et al. Antibodies against cyclic citrullinated peptide
and IgA rheumatoid factor predict the development of rheumatoid arthritis.
Arthritis Rheum. 48, 2741–2749 (2003).
11. Huizinga, T. W. et al. Refining the complex rheumatoid arthritis phenotype
based on specificity of the HLA-DRB1 shared epitope for antibodies to
citrullinated proteins. Arthritis Rheum. 52, 3433–3438 (2005).
12. Verpoort, K. N. et al. Association of HLA-DR3 with anti-cyclic citrullinated
peptide antibody-negative rheumatoid arthritis. Arthritis Rheum. 52,
3058–3062 (2005).
13. Willemen, A., Trouw, L. A., Toes, R. E. & Huizinga, T. W. The influence of
ACPA status and characteristics on the course of RA. Nat. Rev. Rheumatol. 8,
144–152 (2012).
14. Stastny, P. Mixed lymphocyte cultures in rheumatoid arthritis. J. Clin. Invest.
57, 1148–1157 (1976).
15. Raychaudhuri, S. et al. Five amino acids in three HLA proteins explain most of
the association between MHC and seropositive rheumatoid arthritis. Nat.
Genet. 44, 291–296 (2012).
16. Shadick, N. A. et al. Opposing effects of the D70 mutation and the shared
epitope in HLA-DR4 on disease activity and certain disease phenotypes in
rheumatoid arthritis. Ann. Rheum. Dis. 66, 1497–1502 (2007).
17. van der Woude, D. et al. Protection against anti-citrullinated protein antibody-
positive rheumatoid arthritis is predominantly shared with HLA-
DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated
protein-antibody-positive and anti-citrullinated protein antibody-negative
rheumatoid arthritis in four European populations. Arthritis Rheum. 62,
1236–1245 (2010).
18. Feitosa, A. L. et al. Protective effect of noninherited maternal HLA-DR
alleles on rheumatoid arthritis development. Proc. Natl Acad. Sci. USA 104,
19966–19970 (2007).
19. Chiz, R. M. et al. Predominantly naturally processed peptides bound to
HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size.
Nature 358, 784–788 (1992).
20. Ceyhudo, J. A. et al. Composition of the HLA-DR-associated human thymus
peptidome. Eur. J. Immunol. 43, 2273–2282 (2013).
21. Adamopoulou, E. et al. Exploring the MHC-peptide-matrix of central tolerance
in the human thymus. Nat. Commun. 4, 2039 (2013).
22. Sniders, A. et al. An HLA-DRB1-derived peptide associated with protection against rheumatoid arthritis is naturally processed by human APCs. J. Immunol. 166, 4997–5003 (2001).
23. Romero, V. et al. Immune-mediated pore-forming pathways induce cellular hypercitrullination and generate citrullinated autoantigens in rheumatoid arthritis. Sci. Transl. Med. 5, 209ra150 (2013).
24. di Marzo Veronese, F. et al. Autoimmune cytotoxic T lymphocytes in human immunodeficiency virus type 1-infected subjects. J. Exp. Med. 183, 2509–2516 (1996).
25. Propato, A. et al. Apoptotic cells overexpress vinculin and induce vinculin-specific cytotoxic T cell cross-priming. Nat. Med. 7, 807–813 (2001).
26. Oldstone, M. B., Nerenberg, M., Southern, P., Price, J. & Lewicki, H. Virus infection triggers insulin-dependent diabetes mellitus in a transgenic model: role of anti-self (virus) immune response. Cell 65, 319–331 (1991).
27. Wucherpfennig, K. W. & Strominger, J. L. Molecular mimicry in T-cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell 80, 695–705 (1995).
28. van de Stadt, L. A. et al. Monoclonal anti-citrullinated protein antibodies selected on citrullinated fibrinogen have distinct targets with different cross-reactivity. J. Autoimmun. 52, 351–357 (2013).
29. Gregersen, P. K., Silver, J. & Winchester, R. J. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. 30, 1205–1213 (1987).
30. de Vries, N., Tijssen, H., van Riel, P. L. & van de Putte, L. B. Reshaping the shared epitope hypothesis: HLA-associated risk for rheumatoid arthritis is encoded by amino acid substitutions at positions 67–74 of the HLA DRB1 molecule. Arthritis Rheum. 46, 921–928 (2002).
31. van der Helm-van Mil, A. H. et al. An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility. Arthritis Rheum. 52, 2637–2644 (2005).
32. Shadick, N. A. et al. Opposing effects of the D70 mutation and the shared epitope in HLA-DR4 on disease activity and certain disease phenotypes in rheumatoid arthritis. Ann. Rheum. Dis. 66, 1497–1502 (2007).
33. Hensvold, H. A. et al. Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins. Ann. Rheum. Dis. 74, 375–380 (2013).
34. Enouz, S., Carrie, L., Merkler, D., Bevan, M. J. & Zehn, D. Autoimmune T cells bypass negative selection and respond to self-antigen stimulation during infection. J. Exp. Med. 209, 1769–1779 (2012).
35. Hand, T. W. et al. Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. Science 337, 1553–1556 (2012).
36. Liu, X., Zou, Q., Zeng, R., Fang, Y. & Wei, H. Analysis of pH-dependent properties of proteins predicted using PROPKA. Bioinformatics 16, 1157–1166 (2000).
37. Lee, K. H., Wucherpfennig, K. W. & Wiley, D. C. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. Nat. Immunol. 2, 501–507 (2001).
38. Muthur, V. L. & Stern, L. J. The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of peptide binding. Structure 5, 1385–1396 (1997).
39. Freemont, D. H., Hendrickson, W. A., Marrack, P. & Kappler, J. Structures of an MHC class II molecule with covalently bound single peptides. Science 272, 1008–1016 (1996).
40. Rostkowski, M., Olsson, M. H., Sondergaard, C. R. & Jensen, J. H. Graphical analysis of pH-dependent properties of proteins predicted using PROPka. BMC Struct. Biol. 11, 6 (2011).
41. Simonsson, T. et al. Computational protein design: the Proteus software and selected applications. J. Comput. Chem. 34, 2472–2484 (2013).
42. Brooks, B. R. et al. CHARMM: the biomolecular simulation program. J. Comput. Chem. 30, 1545–1614 (2009).
43. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J. Comput. Chem. 29, 1859–1865 (2008).
44. Mackrell, J. R. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 3586–3616 (1998).
45. Osborn, M. W., Chandrasekhar, J., Madura, J. D., Impy, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935 (1983).
46. Neria, E., Fischer, S. & Karplus, M. Simulation of activation free energies in molecular systems. J. Chem. Phys. 105, 1902–1921 (1996).
47. Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald: an N·log(N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092 (1993).
48. Nose, S. A. Unified formulation of the constant temperature molecular-dynamics methods. J. Chem. Phys. 81, 511–519 (1984).
49. Hoover, W. G. Canonical dynamics - equilibrium phase-space distributions. Phys. Rev. A 31, 1695–1697 (1985).
50. Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. Constant-pressure molecular-dynamics simulation - the Langevin Piston method. J. Chem. Phys. 103, 4613–4621 (1995).
51. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of N-alkanes. J. Comput. Phys. 23, 327–341 (1977).
52. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–46 (1996).
53. Tamamis, P., Morikis, D., Floudas, C. A. & Archontis, G. Species specificity of the complement inhibitor compstatin investigated by all-atom molecular dynamics simulations. Proteins 78, 2655–2667 (2010).
54. Tamamis, P. et al. Insights into the mechanism of CsA5 inhibition by PMX53 via implicit solvent molecular dynamics simulations and docking. BMC Biophys. 7, S 5 (2014).
55. Chen, J., Im, W. & Brooks, III C. L. Generalized Born model with a simple smoothing function. J. Comput. Chem. 24, 1691–1702 (2003).
56. van de Stadt, L. A. et al. Large-scale characterization of natural ligands explains the unique gluten-binding properties of HLA-DQ2. J. Immunol. 180, 3268–3278 (2008).
57. Shi, J. et al. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. Proc. Natl Acad. Sci. USA 104, 13731–13737 (2011).
58. Moscuta, A. et al. Structure of celiac disease-associated HLA-DQ8 and non-associated HLA-DQ9 alleles in complex with two disease-specific epitopes. Int. Immunol. 12, 1157–1166 (2000).
59. Lee, K. H., Wucherpfennig, K. W. & Wiley, D. C. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. Nature 428, 501–507 (2001).
60. Murthy, V. L. & Stern, L. J. The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of peptide binding. Structure 5, 1385–1396 (1997).
61. Freemont, D. H., Hendrickson, W. A., Marrack, P. & Kappler, J. Structures of an MHC class II molecule with covalently bound single peptides. Science 272, 1008–1016 (1996).
62. Rostkowski, M., Olsson, M. H., Sondergaard, C. R. & Jensen, J. H. Graphical analysis of pH-dependent properties of proteins predicted using PROPka. BMC Struct. Biol. 11, 6 (2011).
63. Simonsson, T. et al. Computational protein design: the Proteus software and selected applications. J. Comput. Chem. 34, 2472–2484 (2013).
64. Brooks, B. R. et al. CHARMM: the biomolecular simulation program. J. Comput. Chem. 30, 1545–1614 (2009).