Journal of proteome research

pubs.acs.org/jpr

Article

Deciphering the Structural Enigma of HLA Class-II Binding Peptides for Enhanced Immunoinformatics-based Prediction of Vaccine Epitopes

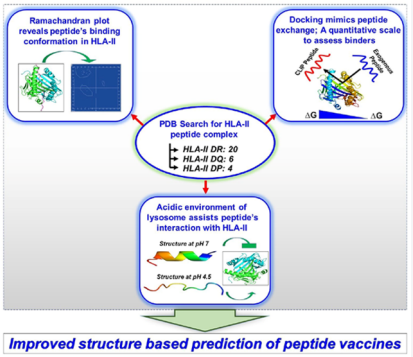
Deepyan Chatterjee, Pragya Priyadarshini, Deepjyoti K. Das, Khurram Mushtaq, Balvinder Singh, and Javed N. Agrewala*

Cite This: *J. Proteome Res.* 2020, 19, 4655–4669

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Vaccines remain the most efficacious means to avoid and eliminate morbid diseases associated with high morbidity and mortality. Clinical trials indicate the gaining impetus of peptide vaccines against diseases for which an effective treatment still remains obscure. CD4 T-cell-based peptide vaccines involve immunization with antigenic determinants from pathogens or neoplastic cells that possess the ability to elicit a robust T helper cell response, which subsequently activates other arms of the immune system. The available *in silico* predictors of human leukocyte antigen II (HLA-II) binding peptides are sequence-based techniques, which ostensibly have balanced sensitivity and specificity. Structural analysis and understanding of the cognate peptide and HLA-II interactions are essential to empirically derive a successful peptide vaccine. However, the availability of structure-based epitope prediction algorithms is inadequate compared with sequence-based prediction methods. The present study is an attempt to understand the structural aspects of HLA-II binders by analyzing the Protein Data Bank (PDB) complexes of pHLA-II. Furthermore, we mimic the peptide exchange mechanism and demonstrate the structural implication of an acidic environment on HLA-II binders. Finally, we discuss a structure-guided approach to decipher potential HLA-II binders within an antigenic protein. This strategy may accurately predict the peptide epitopes and thus aid in designing successful peptide vaccines.



KEYWORDS: *HLA-II binders, peptide HLA-II interaction, torsion angles, structural immunology, molecular docking, molecular dynamics simulation, peptide vaccines, immunoinformatics, structure guided vaccine designing*

1. INTRODUCTION

The CD4 T-cell response is initiated by the presentation of peptides in context with human leukocyte antigen II (HLA-II), expressed on the surface of professional antigen presenting cells (APCs). During the "cut first, bind later" model,[1] HLA-II-mediated antigenic processing and presentation begins by internalization of deleterious proteins, transformed cellular contents, or pathogens by phagocytosis into a cellular organelle, the endosome. The endosome then enters a vesicular pathway to form a lyso-endosomal complex, within which the internalized components are degraded via the acidic and proteolytic activity of lysosomes, which typically have a pH of 4.5 to 5.[2,3] The final stage in the functional maturation of HLA-II begins inside the endosome, which now has multiple antigenic peptides competing to present on the polymorphic antigen binding cleft of HLA-II allomorphs. Through a process of peptide exchange, the epitope exhibiting efficacy is able to displace the endogenous class-II-associated invariant chain peptide (CLIP) to form an HLA-II and peptide complex (pHLA-II) dictated by the hydrogen bonds and noncovalent interactions between them.[4] This complex migrates toward the extracellular membrane to further present and interact with the T-cell receptor (TCR) displayed on the CD4 T cells.[5] Furthermore, the open-ended antigen binding clefts of the heterodimeric HLA-II molecule enable peptides with size of 10- to 25-mer to form the pHLA-II.[6] However, anchoring residues within the binding pocket of the antigen binding cleft typically interact with the 9-mer region of the peptide. Furthermore, they also dictate the binding preference of the TCR-pHLA-II ternary complex.[7,8] Consequently, this region on the peptide is termed as the binding core. Complementing this pathway is the "bind first, cut later" model, wherein epitopes are first scanned within the entire length of the antigenic protein by HLA-II allomorphs. Subsequently, cathepsin cleaves them outside the HLA-II binding groove to generate the free peptides that now bind to the HLA-II allotype before forming the TCR-pHLA-II complex.[1]

Special Issue: Proteomics in Pandemic Disease

Received: June 8, 2020
Published: October 26, 2020

Vaccines have yielded unparalleled success in eliciting robust acquired immunity against infectious and pathological conditions, thus substantially reducing the morbidity and mortality associated with the disease. Classical vaccines constitute the attenuated form of live microbes or complex virulent components.[9,10] Although such formulations are durable and induce long-lasting immunological memory,[11] certain studies indicate probable drawbacks that may be attributed to the superfluous contents of the vaccines, which may instigate reactogenicity such as anaphylaxis, allergic shock, or even autoimmunity.[12] Moreover, producing a stable attenuated strain of a pathogenic organism with an assurance of a lack of reversion to virulence while retaining immunogenic potency is a challenging task.[13]

To overcome such potential problems, peptide vaccines are gaining considerable impetus, as evident by a search of the clinical studies repertoire (https://clinicaltrials.gov/). A total of 182 in phase-I, 115 in phase-II and 7 in phase-III trials are presently underway for peptide vaccine candidates worldwide. Such vaccines aim at inducing long-term prophylactic or therapeutic protection in addition to overcoming the challenges of classical vaccine formulation. These vaccine formulations consist of only defined immunogenic epitopes from a desired antigen that preferentially elicits a protective immune response, and they are devoid of the uncharacterized components of a whole-cell-based vaccine. In addition, virulent reversion or inactivation associated with the classical whole-cell-based vaccines is improbable with synthetic peptide vaccines.[14,15]

A potent cell-mediated immune response can be instigated by peptide vaccines by the HLA-II-mediated processing and presentation of the peptides by the APCs to a TCR expressed on the surface of CD4 T cells. Therefore, an important prerequisite of this approach is the identification of peptide epitopes that show permissive binding, which means, the ability to bind an array of HLA-II allomorphs, and exhibit high immunogenicity. Additionally, such epitopes must not share sequence or structural homology with the host proteins, thus avoiding the chance of generating autoreactivity.[14]

Advancement in the field of immunoinformatics has transformed various aspects of vaccinomics through methods like promiscuous epitope prediction and the *in silico* design of vaccine candidates.[16] Major servers that enable the prediction of promiscuous peptide candidates for HLA-II are based on the linear sequence of the peptide. Although these algorithms are easy to use and fast, some of the key challenges render inconsistencies in their predicted results. These include a distinct possibility of false-positives and false-negatives during prediction and the inability to effectively distinguish self-antigens from non-self-antigens. Furthermore, various prediction servers may also present different results when compared among themselves for the same data set, thus adding to the uncertainty. Thus the algorithms used in designing sequence-based servers may create more confusion than conclusions.[17] Structure-based approaches have been proposed as a potential way to overcome some of the previously described challenges in the field of designing peptide vaccines.[18]

The present study is divided into four parts and is an attempt to understand the structural attributes among the available crystal structures of the pHLA-II complex. The study dialectically presents an enumeration for predicting peptide vaccine candidates that elicit a CD4 T-cell response. We start by statistically evaluating the dihedral angles of peptides within the crystal structure of the pHLA-II complex, analyzing their 9-mer peptide binding core, and assessing the structural contribution of the amino acid physiochemical properties on its secondary structure. Second, we present a quantitative scale for predicting the efficacy of peptides for HLA-II allomorphs by *in silico* mimicking the peptide exchanging mechanism.[4] Furthermore, we demonstrate the seminal contribution of the acidic microenvironment of the lyso-endosome on the structural conformation adopted by HLA-II binding peptides. Finally, we apply the insights gained in evaluating the crystal structures for their ability to distinguish HLA-II binders from nonbinders. This approach aims at addressing the challenges faced in the prediction of CD4 T-cell epitopes by sequence-based *in silico* tools. Subsequently, we discuss an empirically derived workflow that will enable the prediction of accurate epitopes that will significantly enhance the success of peptide-based vaccines.

2. MATERIALS AND METHODS

2.1. Ramachandran Plots and Interpolation Model of HLA-II Epitopes

The dihedral angles, phi (φ) and psi (ψ), of the crystallized HLA-II binding peptides were measured using the Ramachandran window plot of the molecular visualization tool DeepView (Swiss-PDB viewer).[19] Next, using GraphPad Prism version 6.01, an interpolation curve was plotted through the observed values of dihedral angles among amino acid residues in the crystal structure of these HLA-II binding peptides after grouping these peptides according to their size (Table S1). The equation used to interpolate the data points was a second-order polynomial defined by $y = b0 + b1*x + b2*x^2$. The statistical significance of the generated interpolation curve was tested by means of the observed p value, where p values ≤ 0.05 indicated significant deviation of the data set from the model.

2.2. B-Factor Analysis of HLA-II Epitopes

B-factor values of the studied peptides were obtained from the Protein Data Bank (PDB) coordinate files, and visual depictions were plotted using the B-factor coloring option available in PyMOL (https://pymol.org/pymol.html). Next, we grouped the studied peptides according to their size. Subsequently, for each amino acid in a peptide of a particular length, we calculated the normalized B-factor (B_i) value according to the following equation

$$B_i = \frac{b_i}{\sum_{i=1}^{n} b_i} * 100 \begin{cases} n = \text{the peptide } (\therefore \text{ no. of amino acids}) \\ i \rightarrow 1 \\ b_i = \text{B factor of the amino acid at each position } (n) \end{cases}$$

Finally, we computed the mean and standard deviation (SD) of the obtained normalized B-factor values to evaluate them across the different position for all of the peptides.

2.3. Molecular Docking to Assess the Efficacy of HLA-II Epitopes

Peptide exchange involved during the processing and presentation of HLA-II epitopes[4,6] was mimicked by means of molecular docking and the subsequent refinement of the protein−protein docking solution using the online servers PatchDock[20] and FireDock,[21] respectively. First, we identified the amino acid residues residing within the 4 Å region of the peptide binding site on the HLA-II allomorphs using PyMOL. This region was defined as the receptor binding sites of the

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

studied HLA-II allomorphs. Next, the crystal structure of the CLIP peptide, obtained from PDB ID 3QXA, was used as a ligand to perform molecular docking on the defined peptide binding site of the HLA-II molecules. We also performed redocking of the same antigenic cocrystallized peptide on the HLA-II peptide binding site. Both of these docking steps were subsequently followed by molecular refinement to obtain the free-energy change (ΔG) value. Finally, the binding affinity of the best ranked docked complex was compared with that of the docked CLIP, revealed by the negative ΔG for the docked CLIP and cocrystallized antigenic peptides, to interpret their efficacy.

2.4. Two Sample Logo Assessment of the Prevalence of Amino Acids in HLA-II Binders and Nonbinders

MHCBN version 4.0[22] was used to attain an experimentally tested list of 1327 HLA-II binders and 2150 HLA-II nonbinders, respectively. Subsequently, peptides were grouped according to their size, and Two Sample Logo (http://www.twosamplelogo.org/) was used to determine and visualize the propensity of amino acids at different positions on both of the lists.

2.5. *Ab initio* Structure Prediction of Peptides at pH 7 Using PEP-FOLD

The online server PEP-FOLD[23] (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) was used to predict the 3D structure of the peptides. FASTA sequences of the studied peptides were used as an input, and no reference models were provided to PEP-FOLD, and thus the 3D *ab initio* structures of peptides were generated at a neutral pH with the default settings. Cluster representatives from only the top five clusters were selected for the subsequent structural analysis.

2.6. *Ab initio* Prediction of Peptide Structures at pH 4.5 by MD Simulations and Clustering

A linear structure of the studied peptides, whereby the dihedral angles φ and ψ were kept at 180° and 180°, was generated using the AgrusLab molecular modeling tool (http://www.arguslab.com/arguslab.com/ArgusLab.html). Subsequently, these were subjected to a minimization of 10 000 steps, which included the first 5000 steps of steepest descent and the next 5000 steps of conjugate gradient with a restraint weight of 10 kcal/(mol·Å^2) on the backbone atoms of the peptides. Next, the system was gradually heated for 400 ps at a constant volume to 300 K. The equilibration step was performed for 4 ns in the NPT ensemble. Finally, a production run was performed for 100 ns at pH 4.5 with a salt concentration of 0.1 M. The protonation states of residues were maintained according to the pH of the system. All molecular dynamics (MD) simulations were done in an explicit environment using Amber16, and the trajectories were visualized and analyzed using Visual Molecular Dynamics (VMD)[24] and the cpptraj module.[25]

Next, using the cpptraj module of AmberTool 16, we extracted every fifth frame from the 50 000 frames generated during the MD simulation and appended these snapshots in a separate PDB file format. These extracted 10 000 structures were subjected to clustering analysis using the MaxCluster software suite (http://www.sbg.bio.ic.ac.uk/maxcluster/). The Algorithm used for clustering was the nearest-neighbor method,[26] with initial clustering size kept at default, that is, three. The obtained clusters, their size and spread, along with the cluster representative, that is, the centroid structure, are listed in the Table S2. Only the cluster representatives of the top five bins or fewer were used for further structural analysis.

2.7. Comparing the Root-Mean-Square Deviation and Radius of Gyration of the *Ab Initio* Predicted Peptides with Their Crystal Counterparts

Superposition server SuperPose (http://wishart.biology.ualberta.ca/SuperPose/)[27] was used to compute the global backbone root-mean-square deviation (RMSD) between the cluster representatives of the *ab initio* predicted structures and their crystal structure counterparts. Similarly, the radius of gyration (RG) for all of the respective structures was computed using the online version of CRYSOL (https://www.embl-hamburg.de/biosaxs/atsas-online/crysol.php) with default settings.

2.8. Characterization of the Helical Content and Charge on the Peptides at Different pH Values

The freely assessable servers Agadir (http://agadir.crg.es/) and PROTEIN CALCULATOR version 3.4 (http://protcalc.sourceforge.net/) were used to compute the helical content and the net charge on the studied peptides in different pH ranges. Amino acid sequences of the studied peptides were used as the input for both the servers.

2.9. Evaluating the Structure and Efficacy of Experimentally Tested HLA-II Binders and Nonbinders

The PRoteomics IDEntifications (PRIDE) database (https://www.ebi.ac.uk/pride/)[28] was used to identify peptides that are experimentally tested for their ability to bind HLA-II allomorphs. The peptides with lengths of 13-, 14-, and 15-mer in size were chosen as the test peptides to evaluate the present study. Structures of all of the peptides were first modeled by subjecting to MD simulations at pH 4.5 following the previously given protocol. This was followed by clustering to obtain the top five representative structures using MaxCluster and computing the dihedral angles using DeepView (Swiss-PDB viewer).[19] For the structural analysis, the interpolation curve was plotted through the observed values of the dihedral angles at each position across the length of the test peptides. This was then compared with the values of cocrystallized peptides of 13-, 14-, and 15-mer used in the study. To assess the efficacy of the test peptides, docking and refinement were performed using PatchDock[20] and FireDock[21] on the cluster representatives of all of the test peptides onto HLA-II allomorphs (DRB1*0101; PDB ID 1SJE, DPA1*0103 DPB1*0201; PDB ID 4P57, DQB1*0603; PDB ID 1UVQ). No outliers were excluded from the study.

2.10. Statistical Analysis

All statistical analyses was done using GraphPad Prism version 6.01. Graphs were plotted using GraphPad Prism version 6.01 or MS Excel. The statistical significance was computed using the paired parametric t test or the unpaired t test, and the level of significance was determined using p values, where $p > 0.05$ was nonsignificant (NS) and $P \leq 0.05$ was considered significant ($p \leq 0.05$: *; $p \leq 0.01$: **; $p \leq 0.001$: ***).

3. RESULTS

3.1. Interpreting the Structure Relatedness among HLA-II-Binding Peptides

Screening of the cocrystal structures of HLA-II binders from PDB, which had a size of not less than 10-mer, revealed a total of 30 peptides. Of the obtained peptides, 20 were presented on

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

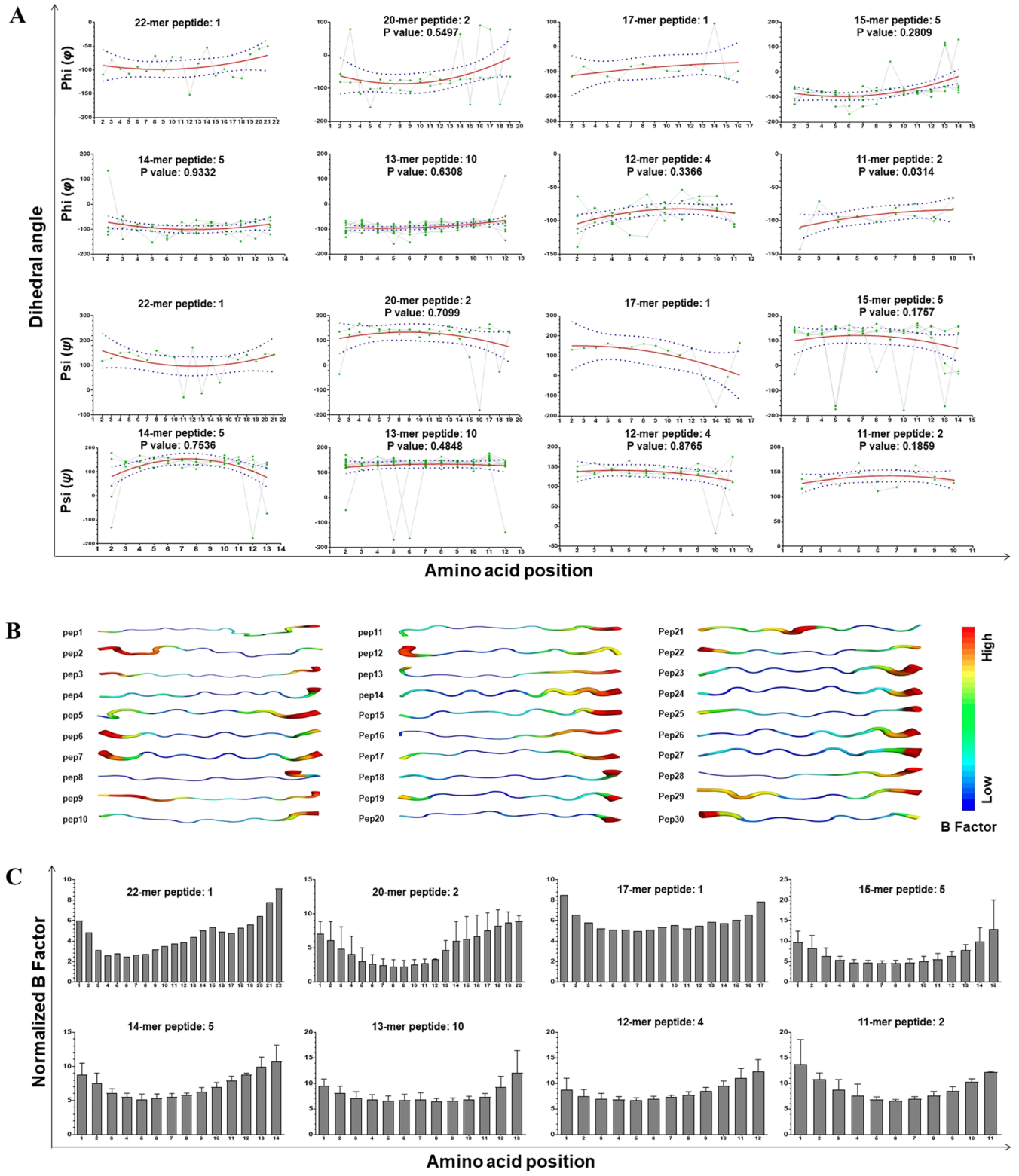


Figure 1. HLA-II bound peptides exhibited structural relatedness across binders of various lengths. The lists of cocrystallized structures of peptides presented on HLA-II were segregated according to their size. (A) The nonsignificant deviation in "p values" for the interpolation curve of the second-order polynomial function plotted across dihedral angles indicates structural relatedness among HLA-II epitopes, with the exception of phi (φ) angles in 11-mer peptides. (B) The color spectrum obtained during B-factor analysis indicates mobility among amino acids in the flanking region, as opposed to core residues. (C) The plotted "mean ± SD" of the normalized B-factor values in amino acids of peptides grouped according to their size corroborates structural conservation among residues in the proximal region. The number of peptides belonging to specified sizes are indicated for panels A and C. In panel A, $p > 0.05$ was considered as nonsignificant.

HLA-II type DR, 4 on HLA-II type DP, and 6 on HLA-II type DQ. Along with the HLA-II type, these were also shortlisted according to the size of the binding peptide; subsequently, dihedral angles of the peptide were computed at each position

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

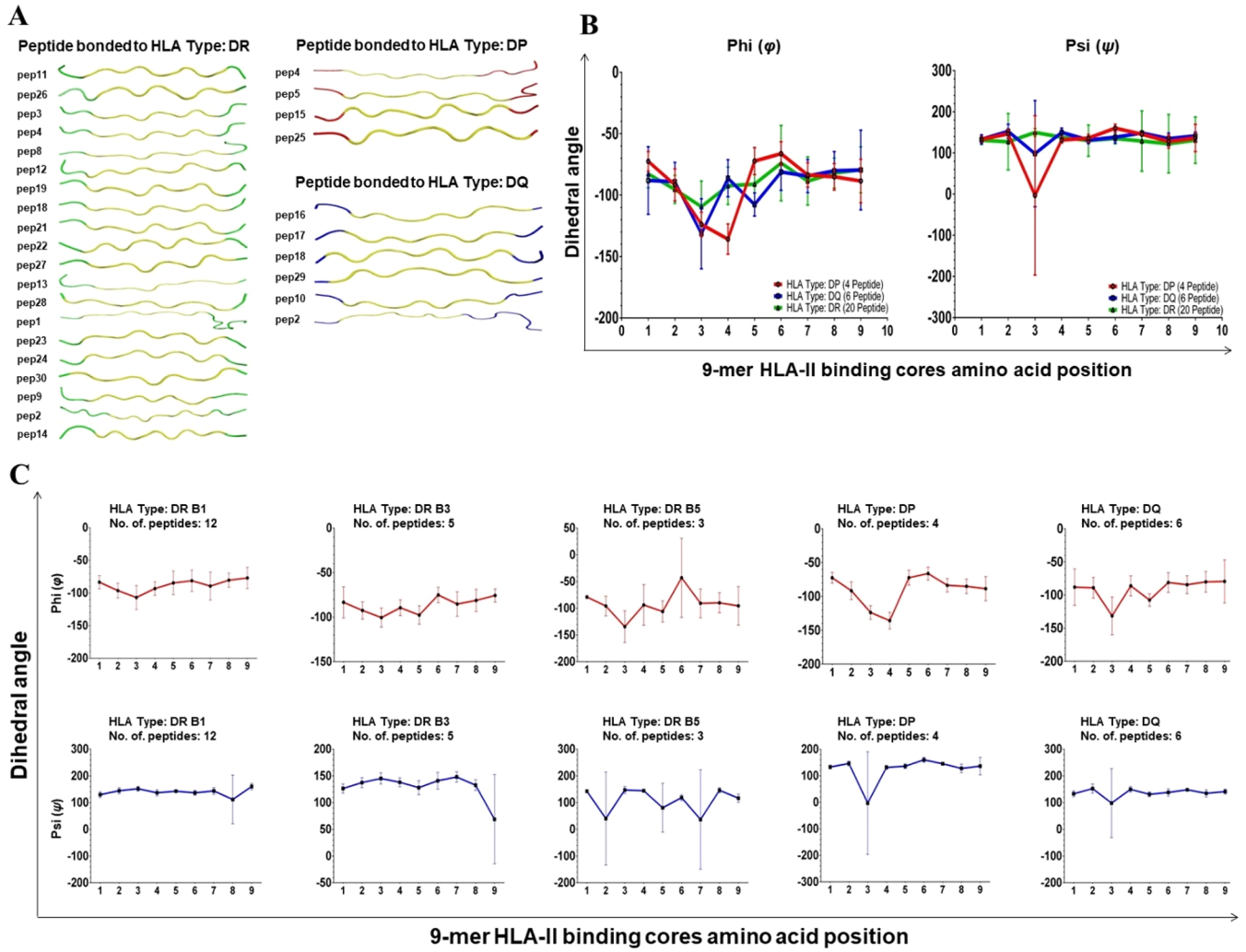


Figure 2. Identifying the 9-mer binding core of peptides presented in association with HLA-II enabled their evaluation, irrespective of their sizes. Anchored residues on the open-ended HLA-II cleft interacted with the 9-mer regions of peptides. (A) 9-mer regions of the studied peptides are highlighted in yellow, and visual analysis reaffirmed the existence of flexible flanking residues. (B) The mean value of the psi (ψ) angles of the studied peptides indicates that HLA-II subtypes exhibited similarity across the entire length, baring position 3, whereas the phi (φ) angles share similar values at positions 2, 7, 8, and 9 of the peptide binding cores. (C) Assertion of the structural similarity in dihedral angles across HLA-II subtypes: DR (B1, B3, B5), DP, and DQ. For panels B and C, structural similarity is established by comparing the mean and standard deviation (SD) of the dihedral angles, φ and ψ, at each position, and the number of peptides used for the computation is also indicated.

and were listed (Table S1, sheets 1−3). For the peptides grouped by size, an interpolation curve of a second-order polynomial (quadratic) function was plotted along the dihedral values, whereby the mean and the SD are depicted as red lines and dotted blue lines in Figure 1A. Instances of extreme values of dihedral angles are isolated and are distinctively outside the interpolation model, indicating the structural semblance among HLA-II binding peptides. This can be statistically verified by the nonsignificant P values of deviation, with the only exception being the φ angles of peptides of 11-mer size. Next, to illustrate the static occupancy and dynamic mobility among amino acids of the HLA-II binding peptide, we visualized their B-factor values by coloring them according to low and high B-factor segments, depicted in blue and red, respectively, in Figure 1B. Furthermore, we normalized the B-factor values of all amino acid residues in the peptides and then plotted the mean and SD at each position along the peptide, grouped according to size in Figure 1C. The color spectrum and heterogeneity among B-factor values in the flanking region, as opposed to the central residues, depict the prevalence of proximal structural conservation among amino acids that may interact with the HLA-II peptide binding core. This observation is in accordance with previous reports.[8] Interestingly, each peptide in the study had a distinct set of amino acids; however, the interpolation curve across the dihedral angles also indicated less SD in the peptide central region, as opposed to the flanking sections (Figure 1A).

3.2. Evaluating the Structural Semblance within the Peptide Core That Interacts with the Anchor Positions on the HLA-II Bed

To further corroborate the existence of structural semblance across the peptide binding core that interacts with the HLA-II bed, we arranged the peptides according to the HLA-II allomorphs, *viz.* DR, DP, and DQ. The 9-mer core binding segments of the bonded peptides were manually identified by their interaction corresponding to the anchor positions of the HLA-II bed. Subsequently, the obtained segments were also cross-verified in the literature[29] and are highlighted in yellow in

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

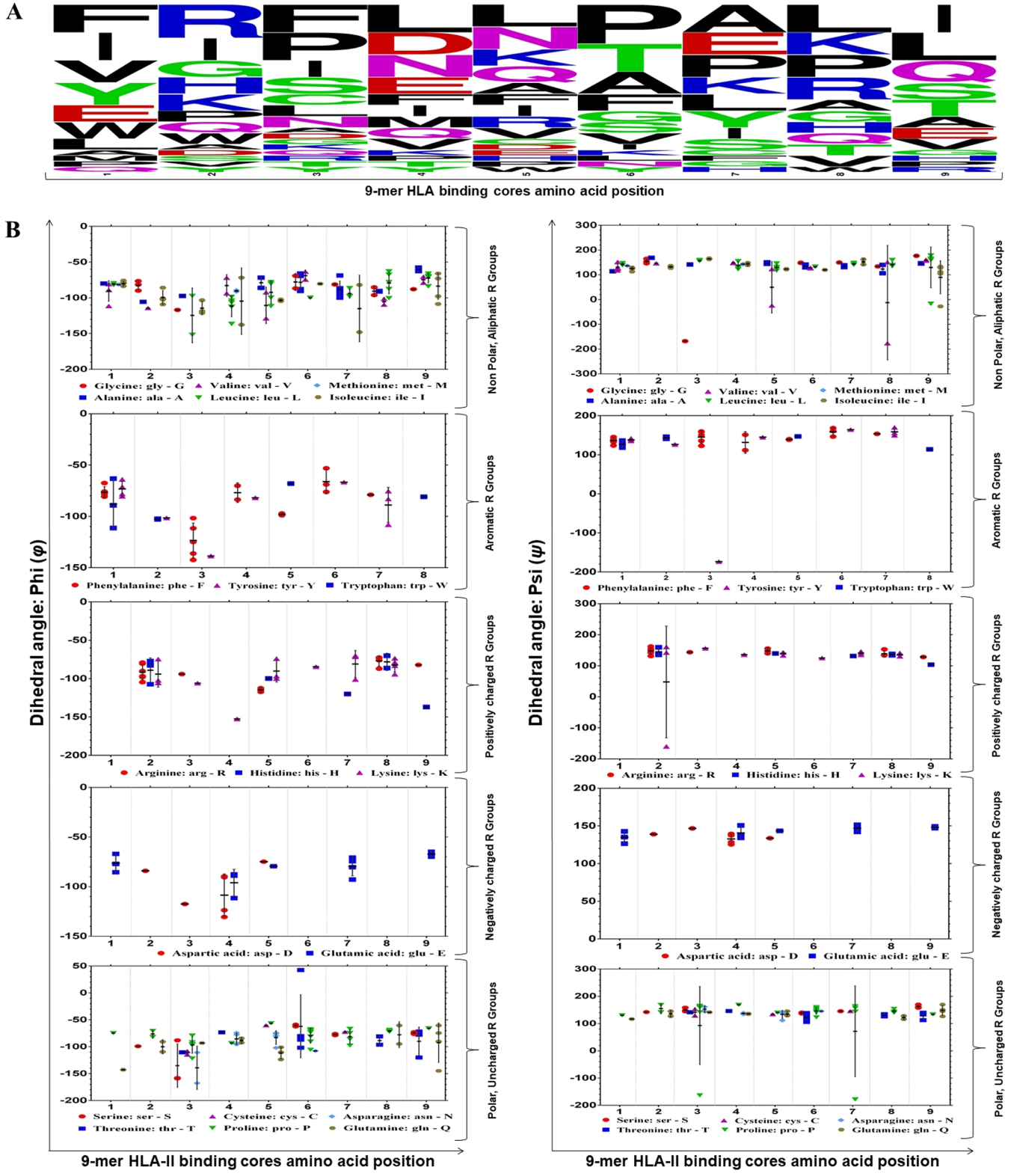


Figure 3. Elucidating the effect on dihedral angles due to the physiochemical properties of amino acid residues within the 9-mer HLA-II binding core. (A) The sequence logo plot depicts the frequency of 20 amino acid types at each position of the 9-mer HLA-II binding core. (B) Computed "mean ± SD" of 20 amino acids grouped by physiochemical properties of the amino acid residues residing in the 9-mer peptide binding core, indicating the contribution of the R group to the behavior of dihedral angles.

Figure 2A. Next, the computed mean and SD of the dihedral angles within the binding core are represented according to all of the HLA allomorphs in Figure 2B and also by the subtypes of HLA-DR, *viz.* DR B1, DR B5, and DR B3, respectively, in Figure 2C. The highly flexible flanking regions of the peptides are attributed to the open-ended cleft of HLA-II,[8] and this can be observed in peptides represented by pep8, pep12, pep13, pep1, pep5, and pep29, as a few examples (Figure 2A).

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

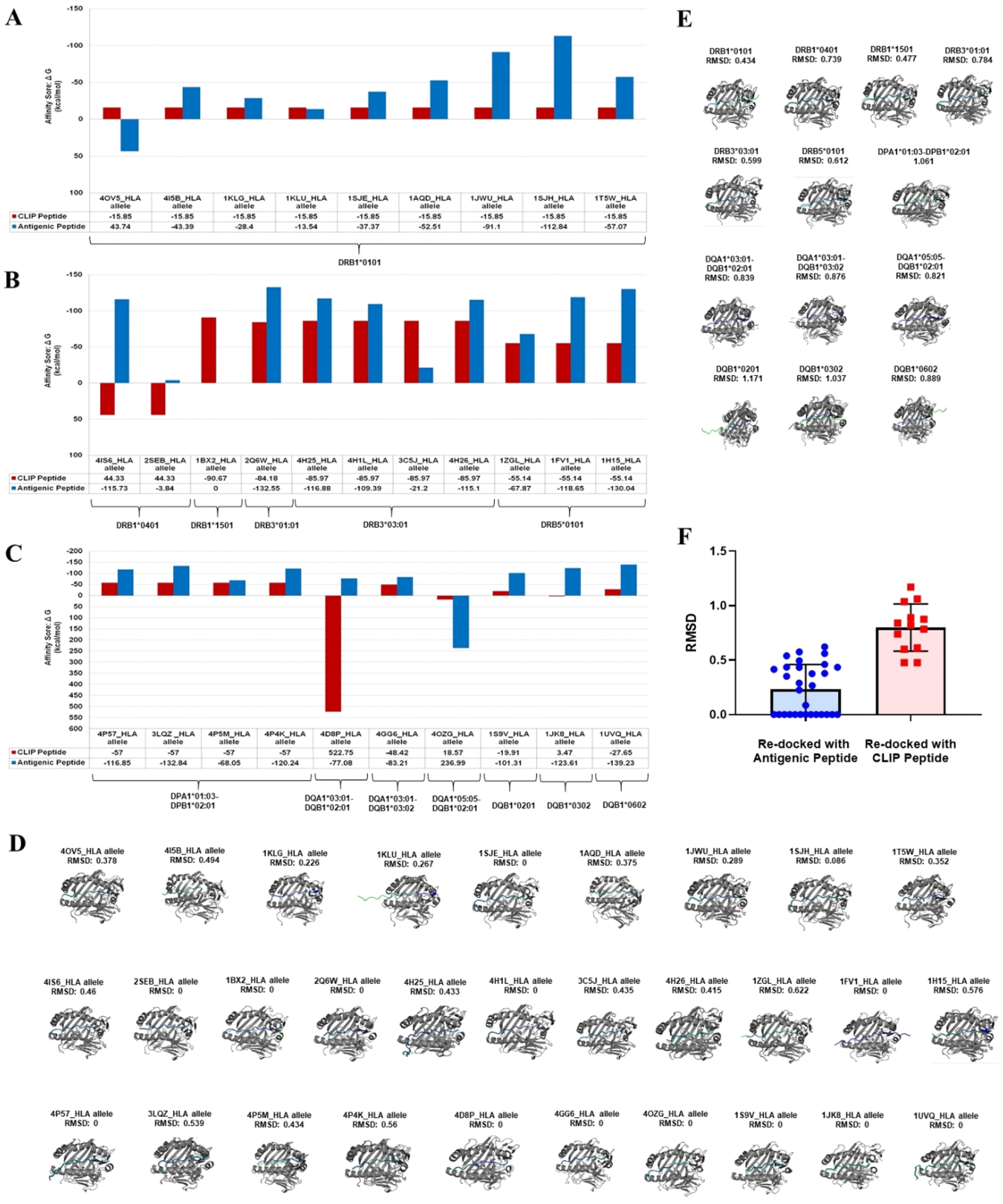


Figure 4. Efficacy of potential HLA-II binders can be evaluated by comparing the docking score with the CLIP peptide. (A−C) 26 times out of 30, antigenic epitopes exhibited stronger efficacy than the CLIP peptide, as observed by the negative free-energy change (ΔG), after docking them to the binding cavity of HLA-II subtypes. (A) Among the exceptions, the peptide represented by PDB ID IKLU is a wild type from the triose-phosphate isomerase (TPI) enzyme and exhibits less efficacy in comparison with CLIP for its HLA-II receptor against its mutant counterpart peptide, represented by PDB ID 1KLG. This is in accordance with the results observed by Sundberg et al. and the tolerance mechanism. RMSDs of redocked peptides are shown in comparison with (D) their crystal counterparts and (E) the CLIP structure obtained from PDB ID 3QXA. Redocked peptides and crystal counterparts are shown in green and blue, respectively. (F) Overall, the low range of RMSD values established the validity of the molecular docking.

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

Furthermore, comparing the 9-mer peptide binding core across all of the studied HLAs enabled us to structurally evaluate them irrespective of their sizes and also nullify the flexible segments within the flanking region of the HLA-II binder. Interestingly, dihedral angle ψ of the core binding regions across all 30 HLA bonded peptides appeared to be more structurally conserved across the entire 9-mer length, except for position 3. However, apart from positions 2, 7, 8, and 9 of the core regions, dihedral angle φ of the HLA-II binders does not appear to share such similarity across the entire 9-mer length (Figure 2B). Furthermore, the high SD of ψ angles at position 3 can be attributed to the peptides presented on the HLA-II allomorphs, DP and DQ, as peptides presented across the three different HLA-II subtypes of DR exhibit a low SD to the computed mean of ψ angles at position 3 (Figure 2C).

3.3. Interpreting the Prevalence of Amino Acid Type and the Structural Effect of Its Physiochemical Properties for the Residing Residues on the 9-mer HLA-II Binding Core

Twenty proteinogenic α-amino acids have structurally distinct side-chain (R group) moieties and thus can be divided according to their physicochemical properties.[30] We evaluated the frequency of different amino acids within the studied 9-mer peptide binding core by plotting a frequency chart of individual amino acids present in it, which are depicted by a sequence logo map according to their position in Figure 3A. Any significant prevalence of a particular type of amino acid at a particular position could not be attributed within the entire region of the core peptide. This may be due to the low data set of only 30 available crystal structures of human HLA-II cocrystallized with peptides that are available in PDB and are used in the present study. Furthermore, to validate the stated observation, we shortlisted an HLA-II peptide data set of known (experimentally tested) binders and nonbinders according to their size from a previous study.[22] Subsequently we plotted the frequency of amino acids at each position, which is depicted in two sample logos in Figure S1. In accordance with the fundamental working principal of sequence-based HLA-II prediction algorithms, we observed the differential preference of specific amino acid types at particular positions among the peptide data sets of known HLA-II binders and nonbinders.

Additionally, in Figure 3B, we evaluated the mean and SD of dihedral angles φ and ψ among the 20 amino acid types, grouped according to five different physiochemical properties ((i) Nonpolar, aliphatic R groups, (ii) aromatic R groups, (iii) positively charged R groups, (iv) negatively charged R groups, and (v) polar, uncharged R Groups), and their position on the HLA-II binding core. Looking back to the high SD of ψ angles at the third position of the core peptide (Figure 2B), we observed that the amino acids glycine, phenylalanine, and proline may occupy positions on the sterically disallowed segments of the Ramachandran plot at this position (Figure 3B). This may be well accredited to fact that glycine has no side chain, so it can occupy dihedral angles on all four quadrants of Ramachandran plot. Similarly, the pyrrolidine ring of proline conformationally restricts it from adopting a helix breaker structure.[31] In essence, comparing the dihedral angles among the amino acids grouped by physiochemical properties enables us to imply the observation that conformationally, ψ angles tend to form a linear shape and are clustered together. However, opposed to ψ angles, the multitude of φ angles in amino acids of the core peptide occupied positions around a much larger region, thus their scattering on the upper left quadrant of the Ramachandran plot (Table S1, sheet 3). Therefore, HLA-II binding peptides are predominantly linear and are devoid of any particular secondary structural component, that is, α helix or β sheet.

3.4. Peptide Exchange between CLIP and Antigenic Epitopes May Be Mimicked via Molecular Docking, Thus Providing a Quantitative Scale to Predict HLA-II Binders

The antigenic presentation pathway across HLA-II molecules involves a peptide exchange mechanism within the compartment of the lyso-endosome, and during this process, only a high-affinity antigenic peptide replaces the HLA-II-associated invariant chain peptide, CLIP.[4] We have successfully mimicked these events using an *in silico* molecular docking approach. In Figure 4, we have compared the binding efficacies of the crystal structure of the CLIP peptide (obtained from PDB ID 3QXA) and the antigenic peptides that we presently study against the appropriate HLA-II receptors. As revealed by the negative ΔG for CLIP and the cocrystallized antigenic peptide, the antigenic peptide has a much stronger affinity as compared with the CLIP peptide in 26 of the 30 available structures of HLA-II subtypes (Figure 4A−C). The HLA-II bonded peptide represented by PDB ID 1KLU was among the four instances where the antigenic peptide exhibits a stronger affinity than the CLIP peptide (Figure 4A). Interestingly both the 1KLU bonded peptide and another peptide represented by PDB ID 1KLG were part of a study by Sundberg et al., which demonstrated the altered potential of wild-type and mutant peptides (represented by PDB IDs 1KLU and 1KLG, respectively) to induce the CD4 T-cell response in the context of a melanoma tumor.[32] Remarkably, in accordance with tolerance mechanisms,[33] the present *in silico* docking simulation was able to account for the inconsequential difference, at the level of sequence and structure (Table S1),[32] between the wild-type and mutant peptides, which can be observed in the altered binding efficacy (ΔG values) for its HLA-II subtype, DRB1*0101 (Figure 4A).

To further validate our observation that docking scores are reflective of peptide−HLA cocrystal structures, RMSD comparisons of the redocked peptide structures with their cocrystallized antigenic counterparts and with the crystal structure of CLIP peptide (PDB ID 3QXA) were performed (Figure 4D,E), respectively. Subsequently, the low range of RMSD values, as shown, endorsed the molecular docking approach as a means to study the peptide exchange mechanism (Figure 4F). Of note, the crystal structure of 3QXA has a CLIP peptide bound to HLA-DR1, and the relatively higher RMSD in comparison with the CLIP peptide docked onto different HLA-II alleles is due the sequence and structural variations among the HLA-II allomorphs (Figure 4E,F).

In essence, we may consider this to be a successful attempt to mimic the peptide exchange mechanism on a quantitative *in silico* scale and also to establish a Goldilocks zone for predicting HLA-II binders and potentially even their immunogenicity. Furthermore, this also has the potential to be used as a scale to quantify the efficacy and promiscuity of the antigenic peptides for HLA-II and thus to assist in designing potential candidates for peptide vaccines.

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

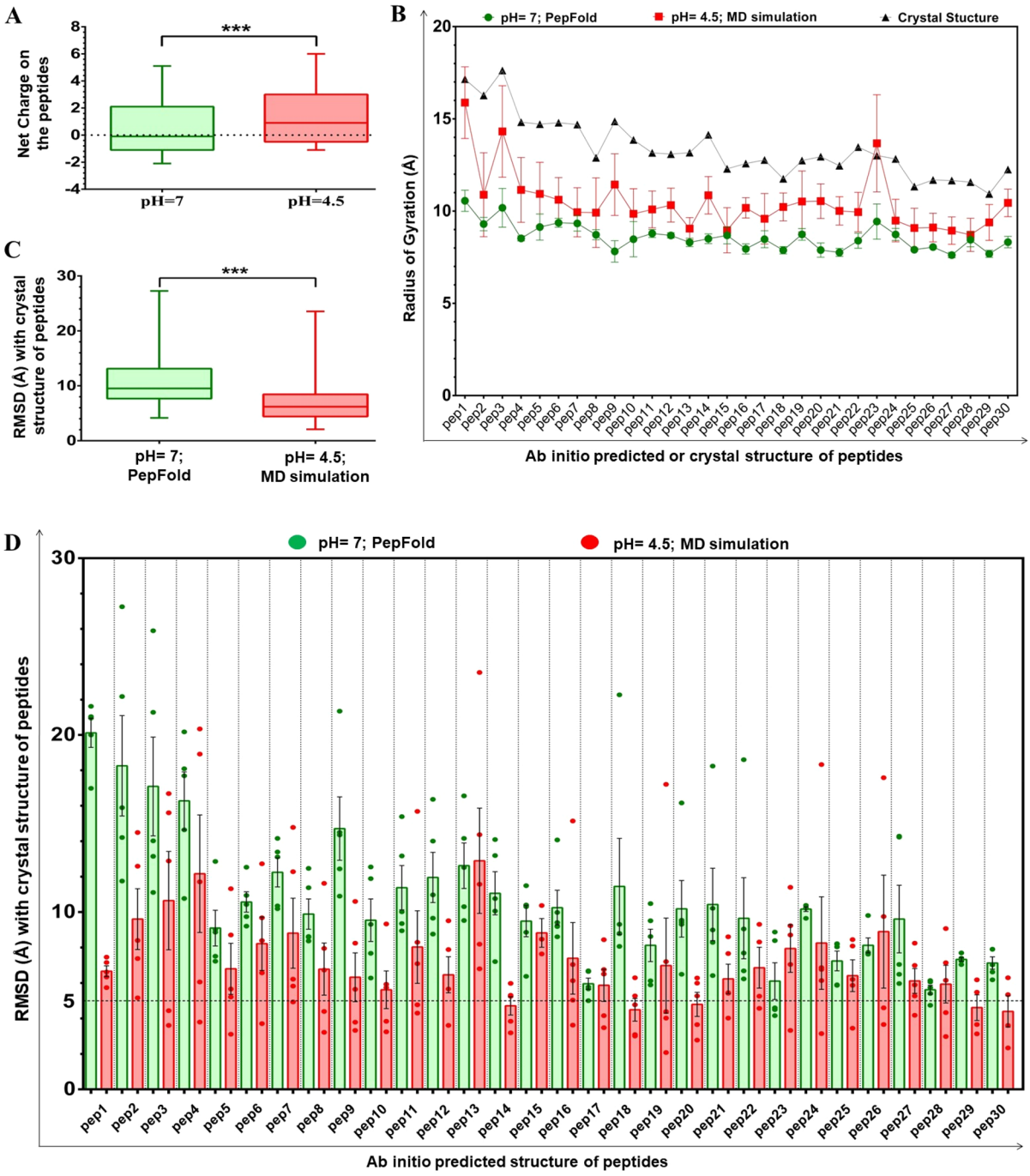


Figure 5. MD simulation suggests that the acidic conditions of the lyso-endosomal compartment helped peptides to adopt a linear conformation, assisting their interaction with the HLA-II. (A) A significant rise in the net charge distribution on the peptides was observed at acidic pH 4.5 compared with neutral pH 7. Cluster representatives of the *ab initio* predicted structures of peptides, obtained at pH 7 (by PEP-FOLD) and pH 4.5 (by MD simulation), were evaluated by crystal structure. (B) A box plot comparing the radii of gyration of peptides suggested that MD simulations at acidic pH also linearize the peptide structures, leading to their resemblance to the crystal structures of the peptides. (C,D) Overall, a significant lessening of structural deviation was observed in MD-simulated peptides at pH 4.5, in contrast with that at neutral pH. Furthermore, in 26 out of 30 peptides, a cluster representative of the MD-simulated peptide at pH 4.5 had an RMSD with crystal structure of <5 Å. Each dot signifies the cluster representatives, and the bar diagram depicts the mean ± SEM. (A,C) Students' paired *t* test and unpaired *t* test were used to determine the significance of deviation in the net charge on the peptides and the RMSD with the crystal structure of peptides. *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$.

3.5. Acidic Conditions within the Lyso-Endosomal Compartment Impart Significance on the Charge and Structure of Antigenic Peptides

Structures of peptides are highly dynamic in nature, and the physicochemical environment can substantially contribute to their secondary structural components.[34,35] Antigenic peptides presented on the HLA-II surface are processed inside the acidic and degradative environment of the lyso-endosomal compartment, typically having a pH of 4.5.[2] We decided to test the structural implication of such an environment on the

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

A

Size	ID	Sequence	Source	Reference
15mer	T_pep1_NB	TIKNSDKNVVLEHFG	Neoantigen from RIF1 Gene	Ott et al., Nature (2017)
	T_pep2_B	AEMVIHHQHVQDCDE	Yelllow fever virus NS5	Chen B et al., Nature Biotechnlogy (2019)
14mer	T_pep3_NB	PVFTHENIQGGGVP	Neoantigen from SORBS 2 Gene	Tran et al., Science (2015)
	T_pep4_B	RLGVRATRKTSERS	Hepatitis C genome polyprotein	Chen B et al., Nature Biotechnlogy (2019)
13mer	T_pep5_NB	DAIEAKKRRQQNF	Neoantigen from STK4 gene	Yossef et al., JCI Insight (2018)
	T_pep6_B	FGQAAAGDKPSLF	Trypanosoma surface antigen-2	Chen B et al., Nature Biotechnlogy (2019)

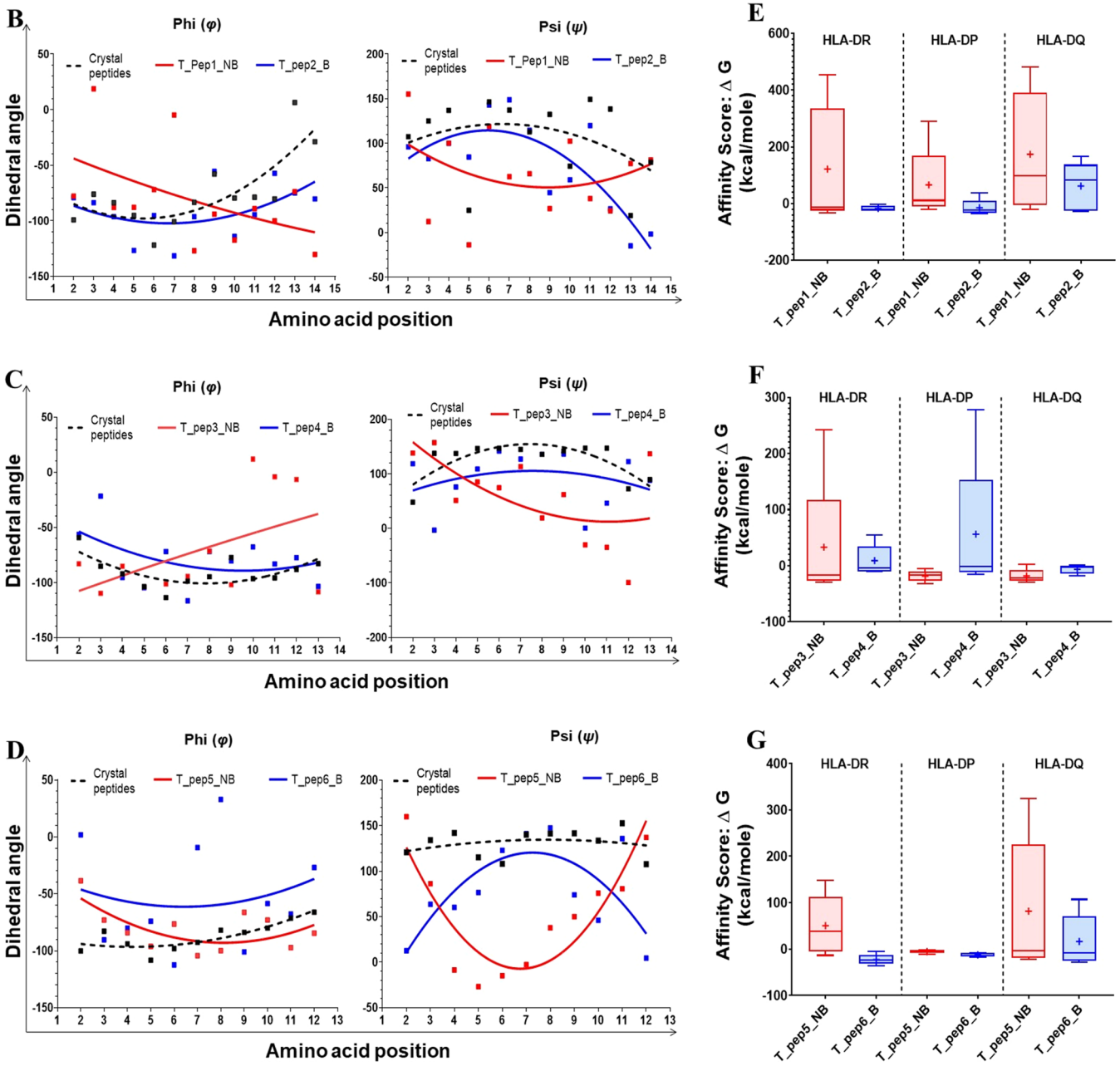


Figure 6. Distinct structure and binding efficacy differentiate the HLA-II binders from nonbinders. (A) A list of test peptides with established immunogenicity were selected from the PRIDE database to differentiate HLA-II nonbinders from binders, depicted in red and blue, respectively. (B–D) Interpolation curves across the dihedral angles of test peptides indicate a dissimilar conformation of nonbinders in comparison with the crystal peptides and HLA-II binding test peptides. (E–G) This was also substantiated in terms of the overall low efficacy of nonbinding test peptides as compared with the HLA-II binding test peptides.

peptides. Amino acids are zwitterions; however, with decreasing pH, they typically become anionic, as the H^+ ion adds to the carboxyl group (COO^-). This can be observed in the studied peptides, that is, a rise in the net charge with the change of pH from neutral to acidic (Figure 5A). A progressive decrease in the helical content of the studied peptides with the decrease in the pH from 8.6 was also noticed (Figure S2A). Therefore, we have predicted the 3D structures of the studied peptides at pH 7 (using PEP-FOLD) and pH 4.5 (using MD simulation), followed by structurally evaluating their cluster

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

representatives at both of the pH values in the context of their crystal structure counterpart (Figure S2B, Table S2). Comparing the average RG values among the top five cluster representatives of the *ab initio* predicted structures reveals distinct structural conformations among the same peptides at different pH values. Interestingly, the average RG of the predicted peptide structures at pH 4.5 was between the RG values of the crystal structure and the predicted peptide structures at pH 7 (Figure 5B). Subsequently, we computed the RMSDs of the cluster representatives of the *ab initio* predicted peptide structures with their crystal structure counterparts. The comparison of the RMSDs presented in Figure 5C demonstrates that the MD simulation under acidic conditions significantly reduced the predicted peptide structural deviation compared with its counterpart at neutral pH 7. Furthermore, 26 times out of 30, at least one or more cluster-representative peptide obtained via post-pH-mediated MD simulation had an RMSD with its crystal counterpart of <5 Å (Figure 5D).

We may conclude that HLA-II bonded peptides are predominantly linear and devoid of any regular secondary structural features (Figure 2B, Table S1).[8] Interestingly, the studied peptides adopted a completely different conformation and were predominantly helical at neutral pH. The MD simulation study suggested that an acidic environment may restrict a peptide's ability to adopt a helical conformation and allow it to remain linear, which may facilitate its interaction with the HLA-II receptor.

Comparing the Experimentally Known HLA-II Binders with Nonbinders Revealed a Distinct Structural Orientation and Binding Efficacy

We applied the insights gained while analyzing the peptide HLA-II cocrystallized complexes to evaluate the experimentally established HLA-II binders from nonbinders. A total of six test peptides of 13-, 14-, and 15-mer sizes were selected from the PRIDE database[28] (Figure 6A). Importantly, the HLA-II binders used (T_pep2_B, T_pep4_B, and T_pep6_B, shown in blue) are critical peptides that were also used to test and prove the accuracy of the latest generation of the HLA-II predictor MARIA.[36] Although test peptides T_pep4_B and T_pep6_B were known HLA-II binders, they were reported as nonbinders by NetMHCIIpan with percentile scores of 30 and 10, respectively (>90% score is considered for binders). The remaining three peptides were randomly selected nonbinders, which are neoantigens that were reported in a negative immunogenicity assay (T_pep1_NB, T_pep3_NB, and T_pep5_NB, shown in red).[37−40] As per the used protocol in the present study, we simulated the peptides at pH 4.5 and performed clustering to obtain five representative model structures of the test peptides. The structural comparison of the interpolated dihedral angles of test peptides against those of the cocrystallized peptides used in the study (Figure 1) is shown in Figure 6B−D and Figure S3. Primarily, post-MD simulation, binder test peptides adopted a structural conformation similar to that of cocrystallized peptide structures, whereas the nonbinding counterparts adopted a distinctively dissimilar conformation. To further contemplate the observation, we checked the efficacy of the test peptides against three HLA-II allomorphs (Figure 6E−G). In seven out of nine instances, that is, 77.7%, the known binder test peptides had better efficacy for HLA-II allomorphs over the known nonbinders. Notably, the two instances where the known nonbinder peptide had better affinity, and efficacy for HLA-DP and DQ, were the 14-mer test peptides, despite the distinctive difference in the dihedral angles (Figure 6C,F). Furthermore, the overall efficacy of the known binder test peptides was lower compared with the CLIP peptide. However, this variation could potentially be attributed to the conformation of the ligand (whether it is of a cocrystal structure for CLIP or an *ab initio* modeled structure for the test peptides) used for docking. We also explored the ability of our methodology to identify three well-characterized and validated immunodominant SARS-CoV-2 CD4 epitopes that bind HLA-DRB1*04:01.[41] The predicted peptide structures obtained after pH-mediated MD simulations were linear, and a stable efficacious binding was observed when they docked onto their cognate HLA-DRB1*04:01, with an average affinity score (ΔG) of around −40 kcal/mol (Figure S3D−G).

In essence, there may be a subtle difference in the structural orientation of HLA-II binding peptides, and the use of pH-mediated MD simulations coupled to docking can be a potential method to differentiate such peptides.

4. DISCUSSION

Vaccines against diseases like HIV, hepatitis C, etc. still remain elusive,[42−44] and many available vaccination strategies for diseases such as tuberculosis, malaria, etc. call for improvements.[45,46] Clinical trials of peptide vaccines corroborate their incipient efficiency as an approach for developing or refining the efficiency of vaccine candidates (https://clinicaltrials.gov/). Interestingly, identifying conserved immunodominant epitopes among serological variants of pathogens may aid in designing vaccine candidates with a broad spectrum protection against the serovars.[47] Furthermore, a desirable single or multifragmented epitope may be limited in inducing not only a cell-mediated immune response but also potent humoral immunity[48] and T−B cell reciprocity.[7,49] Therefore, chemically synthesized single or multiple peptide epitopes, with appropriate adjuvants and excipients, may form a safer alternate to whole-cell vaccines in imparting prophylactic or therapeutic protection.

The emergence of immunoinformatics-based tools and techniques has significantly contributed to the understanding of diverse aspects of immunology, such as the etiology of autoimmunity, tumor immunology, and so on.[50,51] In the relatively young field of vaccinomics,[52] algorithms that make predictions are also gaining a considerable impetus in envisaging immunomodulators[16] and vaccine adjuvant.[53] The fundamental premise of *in silico* prediction in peptide vaccines is to shortlist hitherto unknown epitopes as potential vaccine candidates. However, the majority of CD4 epitope prediction tools are still based on the sequence-based approach. The efficacy of such algorithms is inextricably linked to not just the quantity but also, more importantly, the quality of the data set and the training model employed. Although a few of these algorithms intrepidly claim to be reliable, fast, and simple to use, the limited availability of translatable *in vivo* applications in the literature suggests the disquieting prospect of such sequence-based methods. Interestingly, structure-based epitope prediction has the ability to overcome the demerits of sequence-based methods.[18] Thus the *raison d'etre* for structure-based screening approaches over the available sequence-based algorithms is ideally to set a benchmark for the following criteria: (i) Overlapping fragments must be scanned from an array of antigens to eliminate immunosup-

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

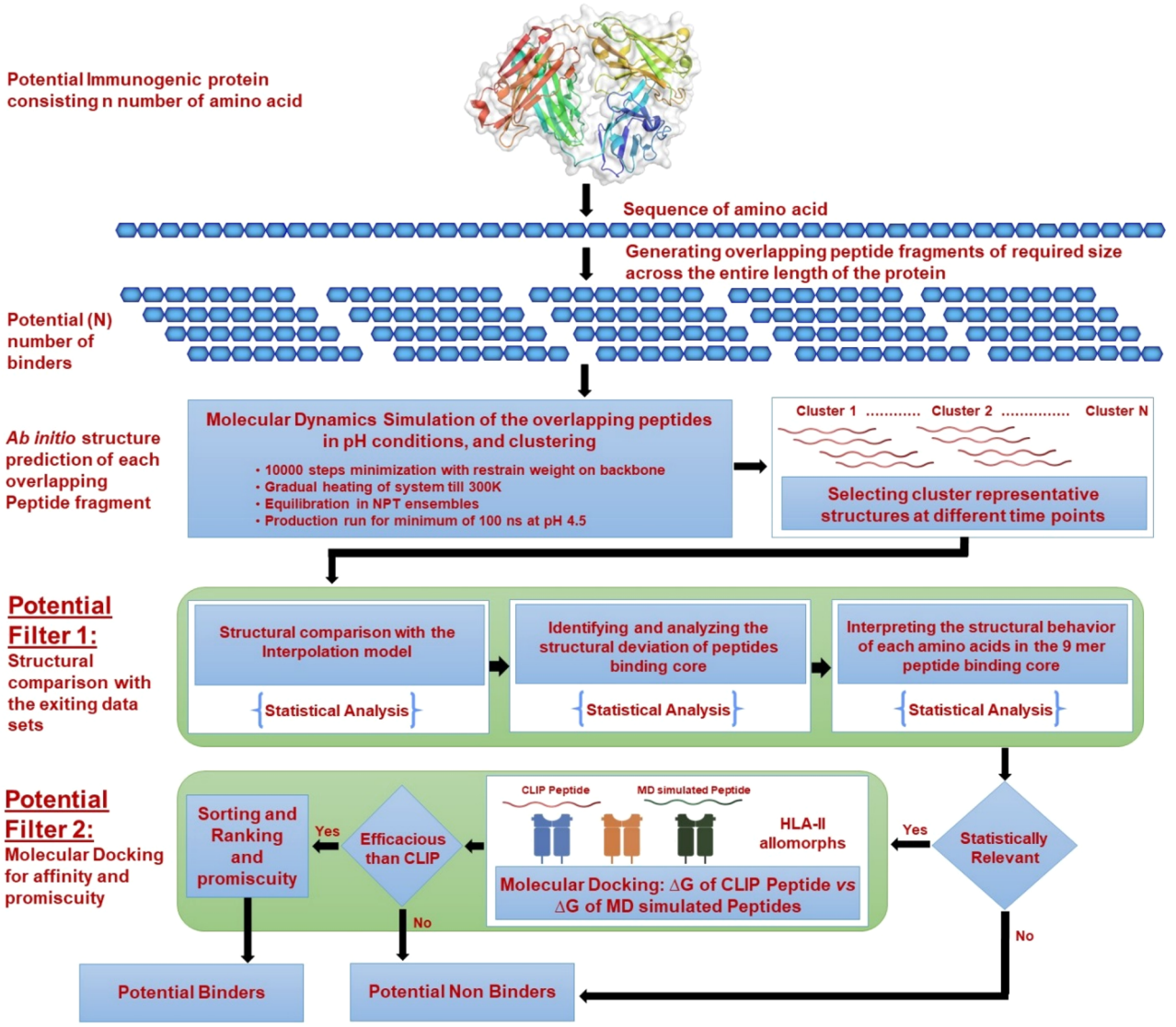


Figure 7. Empiricist layout for deciphering the HLA-II binding epitopes. The scanning process is initiated by generating overlapping peptide fragments of requisite size from the antigenic protein. These peptides can be subjected to MD simulations in a pH environment, followed by clustering to select representative structures. These structures can next be subjected to the first filter, that is, based on their torsion angles, a comparative structural analysis with a known data set to assert the likelihood of interaction with HLA-II. Molecular docking can serve as a second filter to evaluate the efficacy of a potential HLA-II binder and sort them according to their promiscuity.

pressive regions, and a selected number of immunodominant epitopes must be shortlisted. (ii) Such epitopes must not elicit autoreactivity, thus ensuring that they do not exhibit sequence and structural homology with the host proteins. (iii) The promiscuity of such epitopes to transcend the challenge of the extensive HLA-II polymorphism must be considered, ensuring that the result is relevant from a global perspective. (iv) A judicious balance of sensitivity and specificity must be assured at all times, thus eliminating a false sense of reliability while performing the costly wet bench assays of the screened epitopes. Another important value addition in structure-based T-cell epitope prediction could be the subsequent post-differentiation assessment of multimeric pHLA-II and TCR complexes to predict the distinct CD4 T-cell phenotypes that predominantly should include Th1, Th2, and Th17 and induced regulatory T (iTreg) cells.[54] The cytokine milieu during the TCR activation also plays a central role in determining the fate during the effector function.[54,55] Interestingly *in silico* resources also exist to evaluate a peptide's ability to induce cytokines such as IFN-γ, IL-4, IL-10, and IL-17,[56−59] rendering them an essential resource for predicting peptide vaccine candidates.

The present study is an attempt to aid in resolving the conundrums involved the *in silico* prediction of CD4 epitopes,[17] which is essential for designing peptide vaccines. Dihedral angle analysis of the available cocrystallized HLA-II peptide is a potential way to structurally evaluate and distinguish binders from nonbinders. The analysis of the structural aspects of 9-mer core binding segments, along with their amino acid sequences, can form the basis to assess the cognate pHLA-II interaction. We also demonstrate the notable effect of an acidic microenvironment in the lyso-endosome on the peptide structure. The role of acidic, denaturing, and reducing conditions within the degradative environment of lysosomes may not only be limited to imparting a significant structural contribution on the dynamic conformations of a

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

peptide generated during the "cut first, bind later" pathway but also may affect the kinetics of immunogenicity by unfolding antigenic proteins, rendering them accessible to HLA-II allomorphs for the "bind first, cut later" model. Thus for deciphering a potential HLA-II binder, MD simulations at pH 4.5 may be used to predict the starting structure(s) and to obtain the dihedral angles for the peptide conformation. Next, predicted structures can be compared with the dihedral angles of crystal structures of antigenic peptides, illustrated in the interpolation model. Furthermore, a potential HLA-II binder must possess the ability to replace CLIP to present itself on the TCR expressed on CD4 T cells.[4] Our study hitherto demonstrates that molecular docking can be used effectively to obtain the peptide efficacy for an HLA-II receptor. Importantly, this can also be further extended to study the promiscuity of potential HLA-II binders by studying their efficacy for different HLA-II allomorphs. The previously described structure-guided approach can be used as a distinct filter, in combination with statistical tests at each step, for assessing quantitatively and qualitatively the potential HLA-II binders and their promiscuity. The analysis of predictions of the test peptides (known binders and nonbinders) taken from PRIDE databases also corroborates the approach presented in this study. Furthermore, peptides from SARS-CoV-2 have also been shown to be efficient binders using this methodology. A schematic overview, with the stated canonical approach for interpreting HLA-II binders, within antigenic proteins is depicted in Figure 7. Structure-based methods, when implemented in a high-throughput pipeline, may be limited by the requirement of computational power to perform MD simulations, leading to a delay in the availability of results to users. In addition, the limited availability of crystal structures of HLA-II allomorphs with diverse antigenic peptides and the lack of positive and negative controls make it difficult to obtain the accuracy, sensitivity, and specificity of structure-based prediction algorithms. However, these methods are likely to score better than sequence-based methods for predicting potential vaccine epitopes. Thus a separate study employing a hybrid approach involving the structure-based prediction and wet bench validation may provide the requisite standardized data set for feeding into algorithms to be implemented in the *in silico* vaccine prediction techniques. Studies on similar lines have been performed for cancer epitopes and infective diseases;[60−62] however, the data set with the peptide HLA-II binding stoichiometry remains missing.

In conclusion, our study attempts to holistically understand the diverse structural aspects of the cognate peptide and HLA-II interactions, thereby laying down the foundation for a structure-guided screening methodology to discern epitopes for peptide vaccines.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00405.

Figure S1. Position-specific conserved residues observed in the peptide data set of experimentally tested HLA-II binders and nonbinders. Figure S2. Decline in the helical content of peptides with acidic environment. Figure S3. Evaluating the efficacy of the study in deciphering binders and nonbinder (PDF)

Table S1. Computed dihedral angles of the studied peptides, listed according to various criteria that were subsequently used in the present study (XLSX)

Table S2. Clustering information on the studied peptides obtained post-MD simulation in an acid environment (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Javed N. Agrewala − *Immunology Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India; Indian Institute of Technology Ropar, Rupnagar 140001, India;* orcid.org/0000-0003-0227-1042; Email: jagrewala@iitrpr.ac.in

Authors

Deepyan Chatterjee − *Immunology Laboratory and Bioinformatics Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

Pragya Priyadarshini − *Bioinformatics Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

Deepjyoti K. Das − *Immunology Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

Khurram Mushtaq − *Immunology Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

Balvinder Singh − *Bioinformatics Laboratory, CSIR-Institute of Microbial Technology, Chandigarh 160036, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00405

Author Contributions

D.C., B.S., and J.N.A. conceived the project. D.C., D.K.D., and K.M. performed the structural analysis and *in silico* predictions. D.C. and P.P. performed the MD simulation and clustering analysis. D.C., B.S., and J.N.A. analyzed the data. D.C., B.S., and J.N.A. wrote the manuscript. All authors assisted in editing the manuscript and approved it before submission.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by the Department of Biotechnology (BTISnet) and Council of Scientific and Industrial Research (CSIR), India. D.C. and K.M. are recipients of a fellowship from CSIR, India; P.P. of Department of Science and Technology, India; and D.K.D. of Department of Biotechnology, India.

■ REFERENCES

(1) Sercarz, E. E.; Maverakis, E. MHC-guided processing: binding of large antigen fragments. *Nat. Rev. Immunol.* **2003**, *3*, 621−629.

(2) Johnson, D. E.; Ostrowski, P.; Jaumouille, V.; Grinstein, S. The position of lysosomes within the cell determines their luminal pH. *J. Cell Biol.* **2016**, *212*, 677−692.

(3) Roche, P. A.; Furuta, K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat. Rev. Immunol.* **2015**, *15*, 203−216.

(4) Schulze, M. S.; Wucherpfennig, K. W. The mechanism of HLA-DM induced peptide exchange in the MHC class II antigen presentation pathway. *Curr. Opin. Immunol.* **2012**, *24*, 105−111.

(5) Blum, J. S.; Wearsch, P. A.; Cresswell, P. Pathways of antigen processing. *Annu. Rev. Immunol.* **2013**, *31*, 443−473.

(6) Wieczorek, M.; Abualrous, E. T.; Sticht, J.; Alvaro-Benito, M.; Stolzenberg, S.; Noe, F.; Freund, C. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front. Immunol.* **2017**, *8*, 292.
(7) Jorgensen, K. W.; Buus, S.; Nielsen, M. Structural properties of MHC class II ligands, implications for the prediction of MHC class II epitopes. *PLoS One* **2010**, *5*, No. e15877.
(8) Holland, C. J.; Cole, D. K.; Godkin, A. Re-Directing CD4(+) T Cell Responses with the Flanking Residues of MHC Class II-Bound Peptides: The Core is Not Enough. *Front. Immunol.* **2013**, *4*, 172.
(9) Vetter, V.; Denizer, G.; Friedland, L. R.; Krishnan, J.; Shapiro, M. Understanding modern-day vaccines: what you need to know. *Ann. Med.* **2018**, *50*, 110−120.
(10) Pulendran, B.; Ahmed, R. Immunological mechanisms of vaccination. *Nat. Immunol.* **2011**, *12*, 509−517.
(11) Sallusto, F.; Lanzavecchia, A.; Araki, K.; Ahmed, R. From vaccines to memory and back. *Immunity* **2010**, *33*, 451−463.
(12) Skwarczynski, M.; Toth, I. Peptide-based synthetic vaccines. *Chem. Sci.* **2016**, *7*, 842−854.
(13) Fox, G. J.; Orlova, M.; Schurr, E. Tuberculosis in Newborns: The Lessons of the "Lubeck Disaster" (1929−1933). *PLoS Pathog.* **2016**, *12*, No. e1005271.
(14) Purcell, A. W.; McCluskey, J.; Rossjohn, J. More than one reason to rethink the use of peptides in vaccine design. *Nat. Rev. Drug Discovery* **2007**, *6*, 404−414.
(15) Bijker, M. S; Melief, C. J.; Offringa, R.; van der Burg, S. H Design and development of synthetic peptide vaccines: past, present and future. *Expert Rev. Vaccines* **2007**, *6*, 591−603.
(16) Chatterjee, D.; Kaur, G.; Muradia, S.; Singh, B.; Agrewala, J. N. ImmtorLig_DB: repertoire of virtually screened small molecules against immune receptors to bolster host immunity. *Sci. Rep.* **2019**, *9*, 3092.
(17) Gowthaman, U.; Agrewala, J. N. In silico tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion. *J. Proteome Res.* **2008**, *7*, 154−163.
(18) Tong, J. C.; Tan, T. W.; Ranganathan, S. Modeling the structure of bound peptide ligands to major histocompatibility complex. *Protein Sci.* **2004**, *13*, 2523−2532.
(19) Kaplan, W.; Littlejohn, T. G. Swiss-PDB Viewer (Deep View). *Briefings Bioinf.* **2001**, *2*, 195−197.
(20) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363−367.
(21) Mashiach, E.; Schneidman-Duhovny, D.; Andrusier, N.; Nussinov, R.; Wolfson, H. J. FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.* **2008**, *36*, W229−232.
(22) Bhasin, M.; Singh, H.; Raghava, G. P. MHCBN: a comprehensive database of MHC binding and non-binding peptides. *Bioinformatics* **2003**, *19*, 665−666.
(23) Lamiable, A.; Thevenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tuffery, P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* **2016**, *44*, W449−454.
(24) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33.
(25) Roe, D. R.; Cheatham, T. E., 3rd PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(26) Shortle, D.; Simons, K. T.; Baker, D. Clustering of low-energy conformations near the native structures of small proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 11158−11162.
(27) Maiti, R.; Van Domselaar, G. H.; Zhang, H.; Wishart, D. S. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **2004**, *32*, W590−594.
(28) Perez-Riverol, Y.; Csordas, A.; Bai, J.; et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442−D450.
(29) Jensen, K. K.; Andreatta, M.; Marcatili, P.; et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* **2018**, *154*, 394−406.
(30) Janin, J.; Wodak, S.; Levitt, M.; Maigret, B. Conformation of amino acid side-chains in proteins. *J. Mol. Biol.* **1978**, *125*, 357−386.
(31) Gunasekaran, K.; Nagarajaram, H. A.; Ramakrishnan, C.; Balaram, P. Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. *J. Mol. Biol.* **1998**, *275*, 917−932.
(32) Sundberg, E. J.; Sawicki, M. W.; Southwood, S.; Andersen, P. S.; Sette, A.; Mariuzza, R. A. Minor structural changes in a mutated human melanoma antigen correspond to dramatically enhanced stimulation of a CD4+ tumor-infiltrating lymphocyte line. *J. Mol. Biol.* **2002**, *319*, 449−461.
(33) Xing, Y.; Hogquist, K. A. T-cell tolerance: central and peripheral. *Cold Spring Harbor Perspect. Biol.* **2012**, *4*, a006957.
(34) Karle, I. L. Conformational characteristics of peptides and unanticipated results from crystal structure analyses. *Biopolymers* **1989**, *28*, 1−14.
(35) Karle, I. L. Folding, aggregation and molecular recognition in peptides. *Acta Crystallogr., Sect. B: Struct. Sci.* **1992**, *48* (Pt 4), 341−356.
(36) Chen, B.; Khodadoust, M. S.; Olsson, N.; et al. Predicting HLA class II antigen presentation through integrated deep learning. *Nat. Biotechnol.* **2019**, *37*, 1332−1343.
(37) Racle, J.; Michaux, J.; Rockinger, G. A.; et al. Robust prediction of HLA class II epitopes by deep motif deconvolution of immunopeptidomes. *Nat. Biotechnol.* **2019**, *37*, 1283−1286.
(38) Ott, P. A.; Hu, Z.; Keskin, D. B.; et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* **2017**, *547*, 217−221.
(39) Tran, E.; Ahmadzadeh, M.; Lu, Y. C.; et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* **2015**, *350*, 1387−1390.
(40) Yossef, R.; Tran, E.; Deniger, D. C.; Gros, A.; Pasetto, A.; Parkhurst, M. R.; Gartner, J. J.; Prickett, T. D.; Cafri, G.; Robbins, P. F.; et al. Enhanced detection of neoantigen-reactive T cells targeting unique and shared oncogenes for personalized cancer immunotherapy. *JCI Insight.* **2018**, *3*, 3.
(41) Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R. H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **2020**, *27*, 671−680.
(42) Barre-Sinoussi, F.; Ross, A. L.; Delfraissy, J. F. Past, present and future: 30 years of HIV research. *Nat. Rev. Microbiol.* **2013**, *11*, 877−883.
(43) Baseler, L.; Chertow, D. S.; Johnson, K. M.; Feldmann, H.; Morens, D. M. The Pathogenesis of Ebola Virus Disease. *Annu. Rev. Pathol.: Mech. Dis.* **2017**, *12*, 387−418.
(44) Shoukry, N. H. Hepatitis C Vaccines, Antibodies, and T Cells. *Front. Immunol.* **2018**, *9*, 1480.
(45) Nieuwenhuizen, N. E.; Kaufmann, S. H. E. Next-Generation Vaccines Based on Bacille Calmette-Guerin. *Front. Immunol.* **2018**, *9*, 121.
(46) Cockburn, I. A.; Seder, R. A. Malaria prevention: from immunological concepts to effective vaccines and protective antibodies. *Nat. Immunol.* **2018**, *19*, 1199−1211.
(47) Li, W.; Joshi, M. D.; Singhania, S.; Ramsey, K. H.; Murthy, A. K. Peptide Vaccine: Progress and Challenges. *Vaccines (Basel, Switz.)* **2014**, *2*, 515−536.
(48) Tangye, S. G.; Ma, C. S.; Brink, R.; Deenick, E. K. The good, the bad and the ugly - TFH cells in human health and disease. *Nat. Rev. Immunol.* **2013**, *13*, 412−426.
(49) Dai, Y.; Carayanniotis, K. A.; Eliades, P.; et al. Enhancing or suppressive effects of antibodies on processing of a pathogenic T cell epitope in thyroglobulin. *J. Immunol.* **1999**, *162*, 6987−6992.
(50) Pahari, S.; Chatterjee, D.; Negi, S.; Kaur, J.; Singh, B.; Agrewala, J. N. Morbid Sequences Suggest Molecular Mimicry between

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669

Microbial Peptides and Self-Antigens: A Possibility of Inciting Autoimmunity. *Front. Microbiol.* **2017**, *8*, 1938.
(51) Vidyarthi, A.; Khan, N.; Agnihotri, T.; Negi, S.; Das, D. K.; Aqdas, M.; Chatterjee, D.; Colegio, O. R.; Tewari, M. K.; Agrewala, J. N. TLR-3 Stimulation Skews M2Macrophages to M1 Through IFN-alphabeta Signaling and Restricts Tumor Progression. *Front. Immunol.* **2018**, *9*, 1650.
(52) Bragazzi, N. L.; Gianfredi, V.; Villarini, M.; Rosselli, R.; Nasr, A.; Hussein, A.; Martini, M.; Behzadifar, M. Vaccines Meet Big Data: State-of-the-Art and Future Prospects. From the Classical 3Is ("Isolate-Inactivate-Inject") Vaccinology 1.0 to Vaccinology 3.0, Vaccinomics, and Beyond: A Historical Overview. *Front Public Health.* **2018**, *6*, 62.
(53) Nagpal, G.; Chaudhary, K.; Agrawal, P.; Raghava, G. P. S. Computer-aided prediction of antigen presenting cell modulators for designing peptide-based vaccine adjuvants. *J. Transl. Med.* **2018**, *16*, 181.
(54) Zhu, J.; Yamane, H.; Paul, W. E. Differentiation of effector CD4 T cell populations (*). *Annu. Rev. Immunol.* **2010**, *28*, 445−489.
(55) Leung, S.; Liu, X.; Fang, L.; Chen, X.; Guo, T.; Zhang, J. The cytokine milieu in the interplay of pathogenic Th1/Th17 cells and regulatory T cells in autoimmune disease. *Cell. Mol. Immunol.* **2010**, *7*, 182−189.
(56) Dhanda, S. K.; Vir, P.; Raghava, G. P. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct* **2013**, *8*, 30.
(57) Dhanda, S. K.; Gupta, S.; Vir, P.; Raghava, G. P. Prediction of IL4 inducing peptides. *Clin. Dev. Immunol.* **2013**, *2013*, 263952.
(58) Nagpal, G.; Usmani, S. S.; Dhanda, S. K.; Kaur, H.; Singh, S.; Sharma, M.; Raghava, G. P. S. Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci. Rep.* **2017**, *7*, 42851.
(59) Gupta, S.; Mittal, P.; Madhu, M. K.; Sharma, V. K. IL17eScan: A Tool for the Identification of Peptides Inducing IL-17 Response. *Front. Immunol.* **2017**, *8*, 1430.
(60) Abelin, J. G.; Harjanto, D.; Malloy, M.; et al. Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. *Immunity* **2019**, *51*, 766−779.
(61) Pedersen, S. R.; Christensen, J. P.; Buus, S.; et al. Immunogenicity of HLA Class I and II Double Restricted Influenza A-Derived Peptides. *PLoS One* **2016**, *11*, No. e0145629.
(62) Hamrouni, S.; Bras-Goncalves, R.; Kidar, A.; et al. Design of multi-epitope peptides containing HLA class-I and class-II-restricted epitopes derived from immunogenic Leishmania proteins, and evaluation of CD4+ and CD8+ T cell responses induced in cured cutaneous leishmaniasis subjects. *PLoS Neglected Trop. Dis.* **2020**, *14*, No. e0008093.

https://dx.doi.org/10.1021/acs.jproteome.0c00405
J. Proteome Res. 2020, 19, 4655−4669