



Universiteit
Leiden
The Netherlands

Assessment of novel proteins triggering celiac disease via docking-based approach

Atanasova, M.; Dimitrov, I.; Fernandez, A.; Moreno, J.; Koning, F.; Doytchinova, I.

Citation

Atanasova, M., Dimitrov, I., Fernandez, A., Moreno, J., Koning, F., & Doytchinova, I. (2024). Assessment of novel proteins triggering celiac disease via docking-based approach. *Molecules*, 29(1). doi:10.3390/molecules29010138

Version: Publisher's Version



License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3720989>

Note: To cite this publication please use the final published version (if applicable).

Article

Assessment of Novel Proteins Triggering Celiac Disease via Docking-Based Approach

Mariyana Atanasova ^{1,*}, Ivan Dimitrov ¹, Antonio Fernandez ², Javier Moreno ³, Frits Koning ⁴
and Irini Doytchinova ¹

¹ Faculty of Pharmacy, Medical University of Sofia, 1000 Sofia, Bulgaria; idimitrov@pharmfac.mu-sofia.bg (I.D.); idoytchinova@pharmfac.mu-sofia.bg (I.D.)

² European Food Safety Authority, 43126 Parma, Italy; antonio.fernandezdumont@efsa.europa.eu

³ Instituto de Investigación en Ciencias de la Alimentación (CIAL), Consejo Superior de Investigaciones Científicas-Universidad Autónoma de Madrid (CSIC-UAM), Campus of International Excellence—CEI (UAM+CSIC), Nicolás Cabrera, 9, 28049 Madrid, Spain; javier.moreno@csic.es

⁴ Department of Immunohematology and Blood Transfusion, Leiden University Medical Centre, 2333 ZA Leiden, The Netherlands; f.koning@lumc.nl

* Correspondence: matanasova@pharmfac.mu-sofia.bg

Abstract: Human leukocyte antigens (HLAs) are pivotal in antigen processing, presenting to CD4+ T cells, and are linked to autoimmune disease susceptibility. In celiac disease, HLA-DQ2.5 and HLA-DQ8.1 bind gluten peptides on APCs, some recognized by CD4+ T cells, prompting inflammation and tissue damage. While extensively studied experimentally, these alleles lack comprehensive in silico analysis. To explore peptide–HLA preferences, we used molecular docking on peptide libraries, deriving quantitative matrices (QMs) for evaluating amino acids at nine-residue peptide binding cores. Our findings tie specific residue preferences to peptide backbone conformations. Validating QMs on known binders and non-binders showed strong predictive power (89–94% accuracy). These QMs excel in screening protein libraries, even whole proteomes, notably reducing time and costs for celiac disease risk assessment in novel proteins. This computational approach aligns with European Food Safety Authority guidance, promising efficient screening for potential celiac disease triggers.

Keywords: celiac disease; HLA-DQ2.5; HLA-DQ8.1; molecular docking; quantitative matrix; peptide binding prediction



Citation: Atanasova, M.; Dimitrov, I.; Fernandez, A.; Moreno, J.; Koning, F.; Doytchinova, I. Assessment of Novel Proteins Triggering Celiac Disease via Docking-Based Approach. *Molecules* **2024**, *29*, 138. <https://doi.org/10.3390/molecules29010138>

Academic Editor: Anna Marabotti

Received: 29 November 2023

Revised: 22 December 2023

Accepted: 24 December 2023

Published: 26 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Human leukocyte antigens (HLAs) are major histocompatibility complex (MHC) proteins and are involved in antigen processing. Some of the antigens, synthesized in the cell or endocytosed and cleaved into oligopeptides, bind to a specific MHC protein [1]. Different peptides bind to different MHCs. The peptide–MHC complex is transferred to the cell membrane, where is recognized by specific CD4+ T cells and an immune response is initiated.

HLA proteins are the most polymorphic proteins in higher vertebrates. They are divided into class I and class II. More than 25,000 alleles are known for class I and around 10,000 alleles for class II [2]. The HLA proteins of class II consist of two chains: α and β . The peptide binding site is formed by two antiparallel helices and two beta strands. It is open at both ends and accommodates peptides of different lengths, usually between 11 and 25 amino acids. The binding core itself consists of only nine residues. The rest of the residues flank both ends. The HLA class II proteins are grouped in three loci: DR, DQ, and DP. Some of them are associated with susceptibility to autoimmune diseases. For example, DRB1*04:01 is associated with susceptibility to rheumatoid arthritis [3,4], DRB1*03:01 to systematic lupus erythematosus [5,6], DQ8 (DQA1*0301, DQB1*0302) to type 1 diabetes [7,8], DPB1*0201 to chronic beryllium disease [9], etc.

Among the HLA-DQ proteins, DQ2.5 (DQA*05:01, DQB*02:01) and DQ8.1 (DQA*03:01, DQB*03:02) are associated with susceptibility to celiac disease [10,11]. Celiac disease is an autoimmune disorder triggered by gluten ingestion in genetically predisposed individuals. Gluten proteins undergo partial digestion within the human intestine, yielding numerous gluten peptides. These peptides cross the epithelial layer of the mucosa in the small intestine, reaching the lamina propria. Here, an enzyme known as transglutaminase 2 (TG2 or tTG2) catalyses their deamidation. This alteration increases their affinity to HLA-DQ2.5 and HLA-DQ8.1 proteins found on APCs. Subsequently, some of these peptides, when presented on the surface of APCs, are recognized by CD4+ T cells, thereby initiating inflammatory reactions and resulting in tissue damage [12,13]. Symptoms of celiac disease include diarrhoea, abdominal pain, bloating, and weight loss [14]. If left untreated, it can lead to malabsorption of nutrients and various complications. Diagnosis involves blood tests for specific antibodies and confirmatory intestinal biopsy [15]. The treatment requires a strict lifelong gluten-free diet, eliminating wheat, barley, and rye from the patient's diet. Adherence to this diet helps alleviate symptoms and promotes intestinal healing [16]. DQ2.5 is present in 25% of the world population, while DQ8.1 is present in 14% [17]. Only 3% of the individuals carrying any of these two alleles develop celiac disease [18]. However, among the patients with celiac disease, 80–85% carry only DQ2.5, 5–11% carry only DQ8.1, and 3–10% carry both [10,11].

One of the possible reasons for making these proteins susceptible to celiac disease is the structure of one of the pockets in the binding site, i.e., pocket 9. In DQ2.5 and DQ8.1, the Asp^{57β} is mutated to Ala. In most HLA class II proteins, Asp^{57β} forms two salt bridges with Arg^{76α}. These interactions neutralize the charge in this pocket. The mutation of Asp to Ala makes the pocket positively charged, as Ala is not able to form a salt bridge with Arg. Thus, peptides carrying a negatively charged Glu at p9 fit well into pocket 9 and make stable complexes with DQ2.5 and DQ8.1 [10].

Prior studies on the prediction of peptide binding on both alleles via different *in silico* tools were performed [19–27]. Costantini et al. applied homology modelling to model DQ2 and DQ8, and then the modelled complexes with known epitopes were subject to mild energy minimization [22]. The energy of the complexes was assessed via the docking module of the InsightII package, and intermolecular interactions were analysed [22]. Recently, in another study, molecular docking was applied on 26 known immunotoxic peptides to both DQ alleles using the automatic GRAMM-X online server [25]. In our previous study, we applied flexible peptide sidechain molecular docking to derive models for prediction of peptide binding to different HLA-DQ alleles, including HLA-DQ2.5 [27]. The models were implemented via the server EpiDock (<http://www.ddg-pharmfac.net/epidock/>, accessed on 29 November 2023).

Here, we explore the amino acid preferences in peptides that bind to HLA-DQ2.5 and HLA-DQ8.1 through the application of full flexible binding site molecular docking using combinatorial peptide libraries. Each peptide from the libraries underwent docking into the peptide binding site, and the resulting binding scores were normalized to evaluate the amino acid contributions at each peptide position. These contributions effectively differentiated between preferred and non-preferred residues.

To examine the impact of the peptide backbone on the preferences for specific residues, we simulated the binding modes of two peptides targeting DQ2.5 by molecular dynamics. One of the peptides was the α -gliadin peptide QPFPQPELPYP, characterized by a conformationally rigid backbone due to the presence of five Pro residues. The other peptide was the more flexible non-gliadin binder WIEQEGPEYWD. Our findings reveal that the peptide backbone conformation exerts a significant influence on both the binding mode and the preferences for specific residues at certain positions within the peptide binding core.

2. Results

2.1. Docking-Based Quantitative Matrices (QMs) for Prediction of Peptide Binding to HLA-DQ2.5

The X-ray structure of HLA-DQ2.5 in complex with the rigid α -gliadin peptide Q₀P₁F₂P₃Q₄P₅E₆L₇P₈Y₉P₁₀ (pdb code: 6mfg) [28] was used in the molecular docking calculations and molecular dynamics (MD) simulations. The more flexible non-gliadin binding endecamer W₀L₁E₂Q₃E₄G₅P₆E₇Y₈W₉D₁₀ [29], originating from HLA class I α -chain (UniProt: A0A0G2JL56), was modelled by homology using the 6mfg complex as a starting conformation.

The bound peptide in each complex was used as a parent peptide for the construction of a combinatorial peptide library by single amino acid substitution (SAAS), i.e., a residue at a given position in the parent peptide was substituted for with the remaining 19 naturally occurring amino acids—one substitution produces one peptide (Figure 1). The peptide binding core includes the residues from p1 to p9. Next to them at both ends are the flanking residues p0 and p10. The library contains 210 endecamers (19 amino acids \times 11 positions + 1 parent peptide).

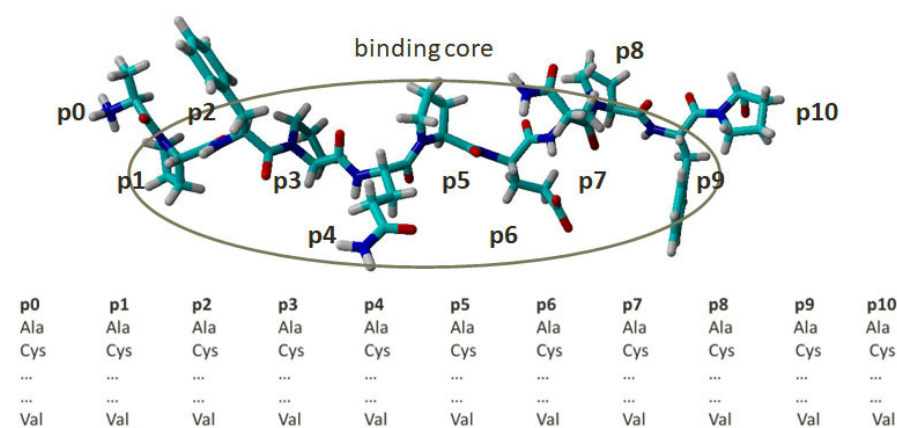


Figure 1. Combinatorial peptide library constructed by single amino acid substitution (SAAS). Positions from 1 to 9 denote the peptide binding core. Positions 0 and 10 are flanking residues. The residue at each position is substituted for by the remaining 19 naturally occurring amino acids. The atoms are coloured by element—carbon in cyan, oxygen in red, nitrogen in blue, and hydrogen in white.

Each peptide from the library was docked into the peptide binding site on HLA-DQ2.5 following the protocol described in Section 4. The protocol was carefully optimized, considering the flexibility of the peptide and binding site, exhaustiveness, and energy range parameters. To generate quantitative matrices (QMs), the binding scores, assigned as affinities, were normalized over both the binding core and the entire peptide. As a result, two QMs were created for each set of parameters, specifically for nonamers and endecamers. The predictive performance of the QMs was evaluated using external test sets, and the QM with the highest *sensitivity*, *specificity*, and *accuracy* was chosen for further analysis.

For the α -gliadin peptide library, the best-performing QM was obtained using the following settings: flexible peptide and binding site, exhaustiveness of 8, energy range of 3 kcal/mol, and normalization over the binding core (Table S1). This QM successfully identified 95% of the binding peptides from the test set and 83% of the non-binders at cutoff 0.1 (Table 1, and Figure S1a). Similarly, for the non-gliadin peptide library, the optimal QM settings were the following: flexible peptide and binding site, exhaustiveness of 8, energy range of 3 kcal/mol, and normalization over the binding core (Table S2). This QM achieved a recognition rate of 93% for the test binders and 89% for the non-binders at cutoff 0.2 (Table 1 and Figure S1a). In both QMs, positive coefficients indicate positive contributions

to the peptides' affinity for the HLA-DQ2.5 protein, while negative coefficients represent negative contributions.

Table 1. Predictive ability of the QMs generated on the peptide-based combinatorial libraries docked on HLA-DQ2.5 and HLA-DQ8.1. *Sensitivity* accounts for the correct prediction of binders, *specificity* for the correct prediction of non-binders, *accuracy* for both.

| QM | <i>Sensitivity</i> | <i>Specificity</i> | <i>Accuracy</i> | <i>Cutoff</i> |
|-----------------------------------|--------------------|--------------------|-----------------|---------------|
| HLA-DQ2.5 | | | | |
| α -gliadin peptide library | 95 | 83 | 89 | 0.1 |
| non-gliadin peptide library | 93 | 89 | 91 | 0.2 |
| HLA-DQ8.1 | | | | |
| α -gliadin peptide library | 94 | 94 | 94 | 0.3 |

Within the peptide binding core, specific anchor positions are responsible for binding to protein pockets with distinct structures and polarity. For HLA-DQ2.5, these anchor positions include p1, p4, p6, and/or p7 and p9 [30]. The preferred amino acids at these anchor positions for the α -gliadin and non-gliadin peptide libraries, as determined by the QMs, are shown in Figure 2. The correlation coefficients range from -0.148 for p6 to 0.711 for p9. This suggests that the preferences for specific residues at the anchor positions are strongly influenced by the backbone conformation.

In the α -gliadin library, the substitution of Pro at p1 with Trp, Phe, Tyr, or Ala enhanced the docking affinity. In the non-gliadin library, replacing Ile with Trp, Ala, Val, Glu, Gln, or Ser led to an increase in predicted affinity. However, the substitution of Ile with Pro at p1 in non-gliadins significantly decreased the affinity. Regarding p4, the preferences in the gliadin library favoured Phe, Trp, Thr, Val, and Glu, while among the non-gliadins, Val, Asp, and Gly were favoured. Pro and Lys were unfavourable in both libraries. The most significant differences between the gliadin and non-gliadin libraries were observed in the preferred and non-preferred amino acids at p6, as indicated by the negative correlation coefficient ($R = -0.148$). Replacing Glu in α -gliadin library with aromatic and aliphatic residues resulted in an increase in the binding scores. Conversely, in the non-gliadin library, substituting Pro for aromatic residues led to a decrease in affinity. At p7 in the gliadin library, a diverse range of amino acids including Trp, His, and Cys were well tolerated, while the presence of Arg was unfavourable. In contrast, non-gliadins showed a preference for hydrophobic amino acids here without any significant disfavour towards specific residues. At p9, replacing Tyr with other aromatic amino acids such as Trp and Phe enhanced the affinity of gliadin peptides. For the non-gliadins, Trp was the only preferred residue at p9, while Pro was the least preferred one.

Peptide–HLA interactions at the anchor positions within the peptide binding site are summarized in Table 2 and visually depicted in Figure S2. The binding pockets are denoted by a capital letter 'P' followed by a number corresponding to the peptide binding position. Trp emerges as a pivotal residue, exerting a positive influence on the affinity of α -gliadin peptides binding to HLA-DQ2.5. Trp was notably favoured at positions p1, p6, p7, and p9, where it established hydrogen bonds with Asn^{62 α} when occupying pockets P6, P7, or P9. Furthermore, Trp was engaged in π - π stacking interactions with Phe^{53 α} , Phe^{11 β} , Phe^{47 β} , and Trp^{61 β} , contributing significantly to peptide binding. Additionally, it participated in hydrophobic interactions with multiple amino acids across both α and β HLA chains. Phe was the prominent residue in pocket P4, forming two hydrogen bonds with Tyr^{9 α} and Asn^{62 α} and demonstrating hydrophobic interactions with Phe^{11 β} , Leu^{26 β} , and Lys^{71 β} . Among the non-preferred residues, Met in pockets P1 and P6 had a detrimental effect on binding, while Pro in pocket P4, Arg in pocket P7, and Lys in pocket P9 also yielded unfavourable outcomes. Arg from p7 formed an intramolecular hydrogen bond with Gln at p6, evidently disturbing the binding. The negative effect of Lys at p9 can be attributed to the presence of the positively charged Arg^{76 α} in pocket P9 [10].

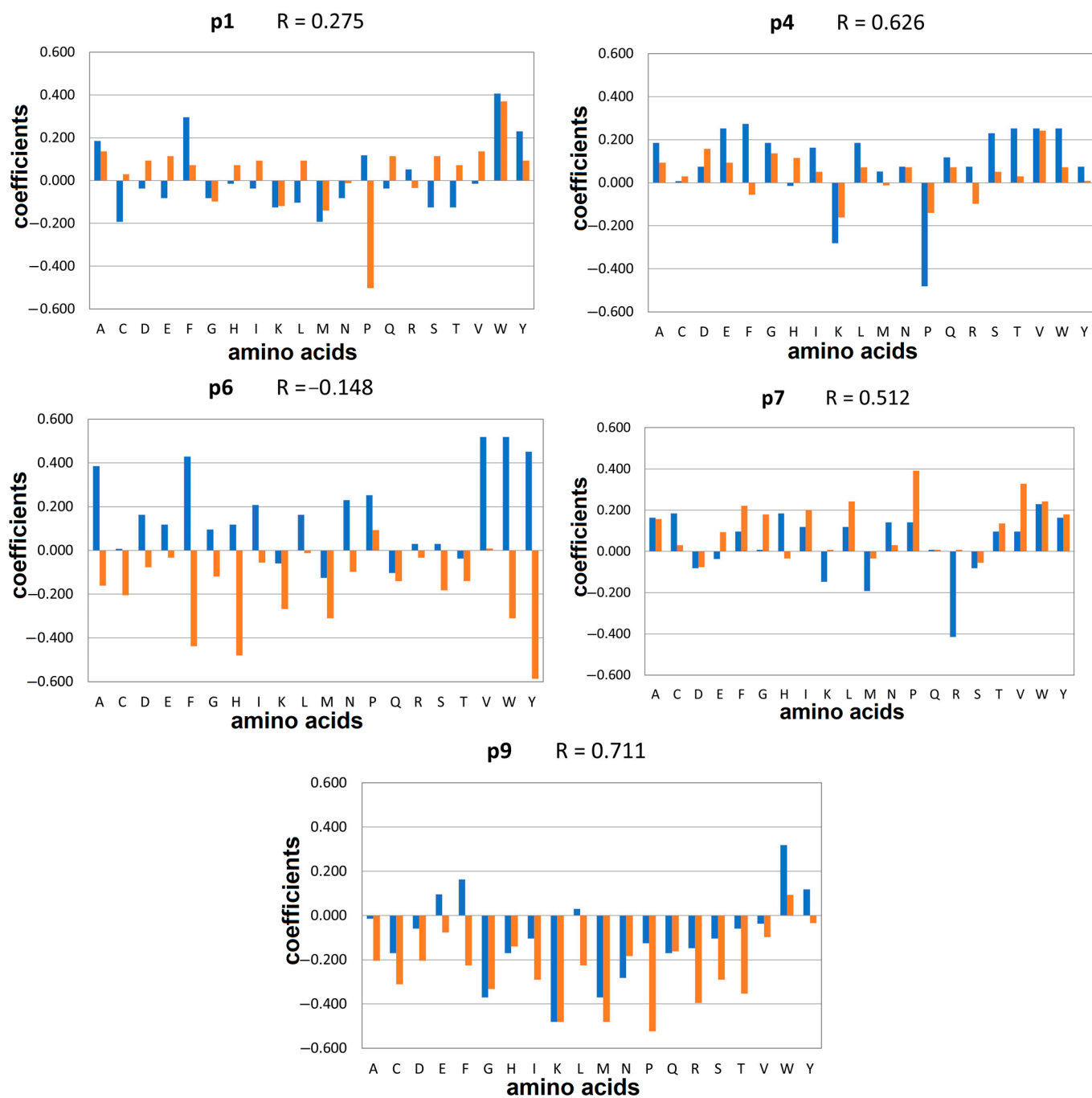


Figure 2. Contributions of the amino acids at the anchor positions p1, p4, p6, p7, and p9 according to the QMs' coefficients for α -gliadin (blue bars) and non-gliadin peptide libraries (orange bars) for HLA-DQ2.5. The correlation coefficients between the contributions are given in the chart titles. Positive coefficients indicate positive contributions and negative coefficients indicate negative contributions.

Table 2. Intermolecular interactions between the most positively/negatively contributing peptide residues from α -gliadin and non-gliadin peptides binding into the corresponding pockets of HLA-DQ2.5 and HLA-DQ8.1. The residues of the corresponding peptide position are given in bold; hb stands for hydrogen bond, hpho for hydrophobic interaction; π - π - π - π contact.

| Pocket | Contribution | DQ2.5 | | DQ8.1 |
|--------|--------------|---|---|---|
| | | α -Glia aa | non-Glia aa | α -Glia aa |
| P1 | positive | W : π - π -F53 α ; hpho—W43 α , F51 α , F54 α , N82 β , L85 β | W : hb—R53 α ; π - π -F33 α ; F52 α , F54 α ; hpho—F33 α , F52 α , F54 α , N82 β , E86 β | F : hb—R53 α ; hpho—Y9 α , F54 α , L85 β |
| | negative | M : hpho—W43 α , L85 β | P : hpho—F54 α | P : hpho—F54 α |
| P4 | positive | F : hb—Y9 α , N62 α ; hpho—F11 β , L26 β , K71 β | V : hb—Y9 α ; hpho—L26 β , K71 β , V78 β | F : hb—Y9 α , N62 α ; hpho—F11 β , L26 β , T28 β , K71 β , E74 β |
| | negative | P : hpho—V78 β | K : hb—Y9 α , G10 α ; intramolecular hb with G5; hpho—L26 β , K71 β | Y : hb—Y9 α , N62 α ; hpho—F11 β , T28 β , E74 β , V78 β |
| P6 | positive | W : hb—N62 α ; π - π -F11 β , W61 β hpho—V65 α , L66 α , F11 β , S30 β , W61 β | P : hpho—V65 α , L66 α , F11 β | P : hpho—V65 α , F11 β , Y30 β |
| | negative | M : hb—N62 α ; hpho—N62 α , V65 α , F11 β | Y : hb—N62 α ; π - π -F11 β , W61 β hpho—V65 α , Y9 β , F11 β | Y : π - π -F11 β hpho—N62 α , V65 α , F11 β |
| P7 | positive | W : hb—N69 α ; π - π -F11 β , F47 β , W61 β ; hpho—F11 β , S28 β , F47 β , W61 β , K71 β | P : hb—N69 α ; hpho—W61 β , K71 β | - |
| | negative | R : hb—N69 α ; intramolecular hb with Q6; hpho—F47 β , W61 β | D : hb—N69 α , K71 β ; hpho—W61 β | - |
| P9 | positive | W : hb—N69 α ; hpho—S72 α , L73 α , I37 β | W : hb—N69 α ; hpho—L73 α , I37 β , W61 β | N : hb—H68 α , N69 α |
| | negative | K : hb—N69 α ; hpho—S72 α | P : no contacts | P : hb—H68 α ; hpho—W61 β |

In the non-gliadin library, Trp emerged as the most preferred residue in pockets P1 and P9, while Pro was the most favoured in pockets P6 and P7, and Val was the most preferred choice in P4 (Table 2, Figure S2). Notably, Tyr^{9 α} , Arg^{53 α} , and Asn^{69 α} were engaged in hydrogen bonding with the peptide residues at positions p2, p1, p7, and p9, respectively. Additionally, Trp at p1 formed π - π stacking interactions with Phe^{33 α} , Phe^{52 α} , and Phe^{54 α} . Conversely, Pro was the least preferred at positions p1 and p9, Lys was most unfavourable at p4, Tyr was least preferred at p6, and Asp was the least preferred at p7.

2.2. MD Simulations of Peptides Carrying Trp at p10

In the gliadin peptide library, among the non-preferred amino acids for the flanking residues p0 and p10, the highest negative contribution belongs to Trp at p10. The value for Trp10 is -0.649 . In the non-gliadin peptide library, this value is 0.136 . As no information about the detrimental contribution of Trp at p10 was available in the literature, we examined the effect of Trp10 on the stability of the peptide–HLA complexes by MD simulations on four peptides: α -gliadin peptide QPFPQPELPYP, α -gliadin peptide with P10W mutation QPFPQPELPYW, non-gliadin peptide WIEQEGPEYWD, and non-gliadin peptide with D10W mutation WIEQEGPEYWW. The complexes were simulated for 1000 ns and the root-

mean-square deviations of atomic coordinates over time (rmsds) and the root-mean-square fluctuations (rmsfs) of $C\alpha$ were recorded at each peptide position per ns. Three runs were performed for each complex. The average rmsd, rmsf, and differences in the fluctuations (Δ rmsfs) between the wild and mutant peptides are given in Figure 3.

The rmsds of the four complexes reached a stationary shape before 100 ns (Figure 3a). Greater deviations were observed for α -gliadin peptide with P10W mutation (Figure 3a). It is evident that the presence of Trp at p10 in the rigid gliadin peptide causes a huge distortion not only on p10 but also on p5, p6, p7, p8, and p9 positions (Figure 3b). The differences in rmsf for p5–p10 were between 1.07 and 2.44 \AA^2 (Figure 3c). No distortion was observed in the more flexible non-gliadin peptide when Trp was present at p10. The differences here were between -0.06 and -0.48\AA^2 .

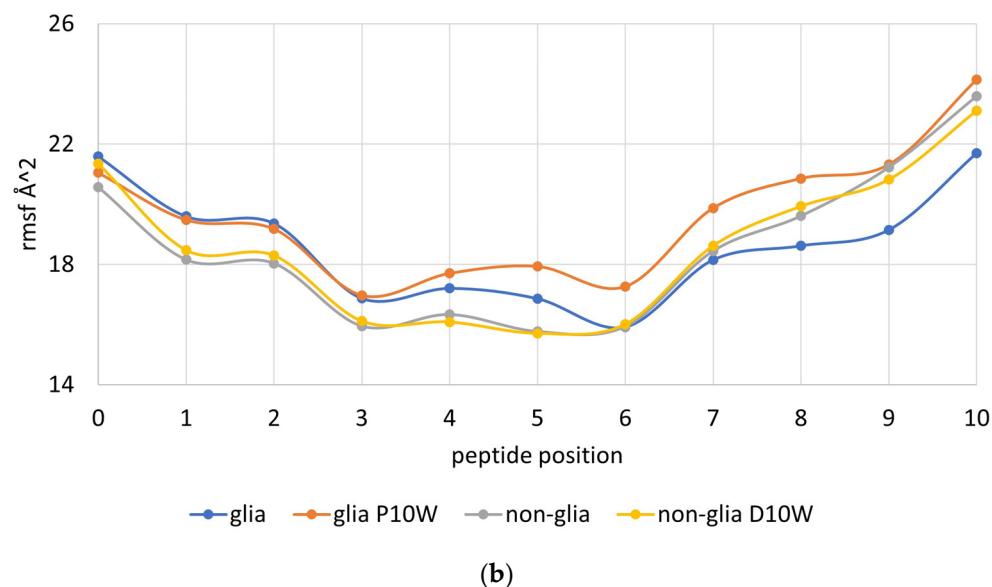
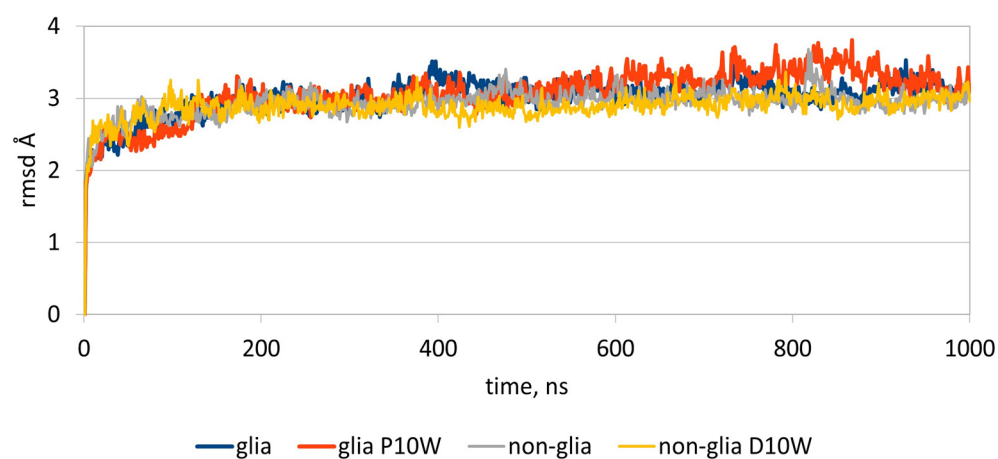


Figure 3. Cont.

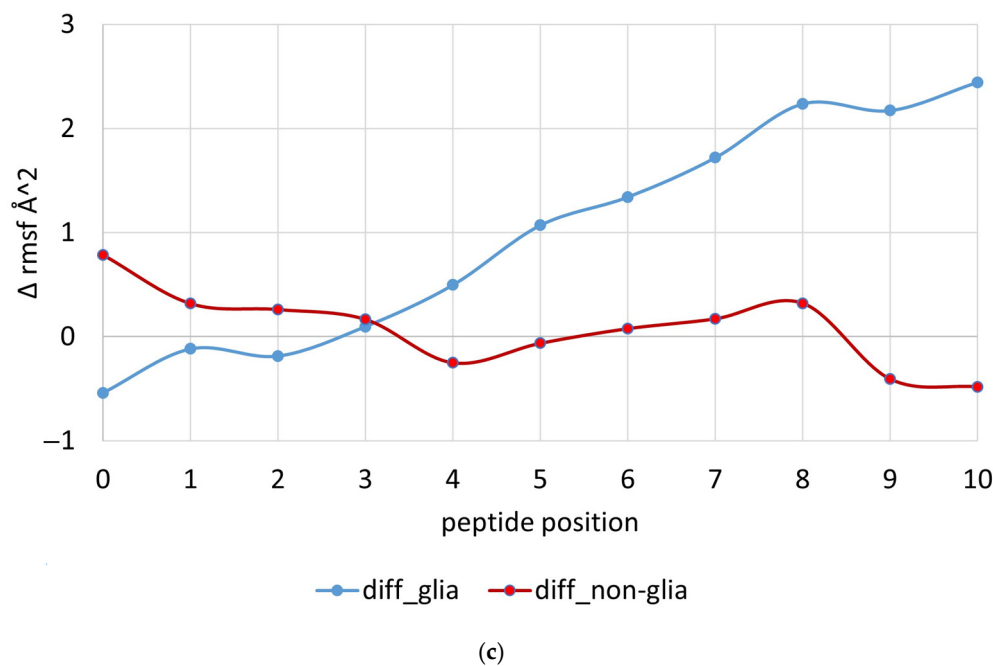


Figure 3. Root-mean-square deviations (rmsds) (a), root-mean-square fluctuations (rmsfs) (b), and differences in the root-mean-square fluctuations (Δ rmsfs) (c) of peptides' C α atoms. In (a,b), α -gliadin peptide (QFPQPPELPYP) is shown in blue, α -gliadin peptide with P10W mutation (QFPQPPELPYW) in orange, non-gliadin peptide (WIEQEGPEYWD) in grey, and non-gliadin peptide with D10W mutation (WIEQEGPEYWW) in yellow. In (c), rmsd differences between the wild and mutant peptides (blue curve) and non-gliadin peptides (red curve) are depicted.

2.3. Docking-Based Quantitative Matrices (QMs) for Prediction of Peptide Binding to HLA-DQ8.1

The X-ray structure of HLA-DQ8.1 in complex with the T-cell epitope from α -gliadin S₋₁G₀E₁G₂S₃F₄Q₅P₆S₇Q₈E₉N₁₀P₁₁ (pdb code: 2nna) [10] served as the basis for the molecular docking studies. The bound peptide parented a combinatorial library consisting of 210 endecamers. Docking simulations were performed using various settings, and the best predictions were obtained with a flexible peptide and protein binding site, an exhaustiveness of 8, an energy range of 3 kcal/mol, and normalization over the binding core (Table S3). The resulting QM achieved a recognition rate of 94% for both binding and non-binding peptides in the external test set at a cutoff of 0.3 (Table 1 and Figure S1b).

Figure 4 displays the bar charts illustrating the preferences and non-preferences of residues at anchor positions p1, p4, p6, and p9 for peptide binding to HLA-DQ8.1 [7]. At p1, Phe and Trp were the most preferred residues establishing hydrogen bonds with Arg^{53 α} (Table 2, Figure S2). Pro was not favoured at this position. At p4, Phe was the most favoured, while Tyr was the least preferred. Both Phe and Tyr were involved in hydrogen bonds with Tyr^{9 α} and Asn^{62 α} . However, the hydroxyl group of Tyr can lead to clashes with Val^{78 β} (Figure S2). At p6, Pro emerged as the most favoured residue, while aromatic residues such as Tyr, Phe, and Trp had detrimental effects, with contributions ranging from -0.7 to -0.6 . For p9, the preferences indicate that Asn, Glu, Thr, and Leu are the favoured residues. For example, Asn formed two hydrogen bonds with His^{68 α} and Asn^{69 α} from pocket P9 (Table 2, Figure S2). Conversely, Pro was the least favoured residue at this position.

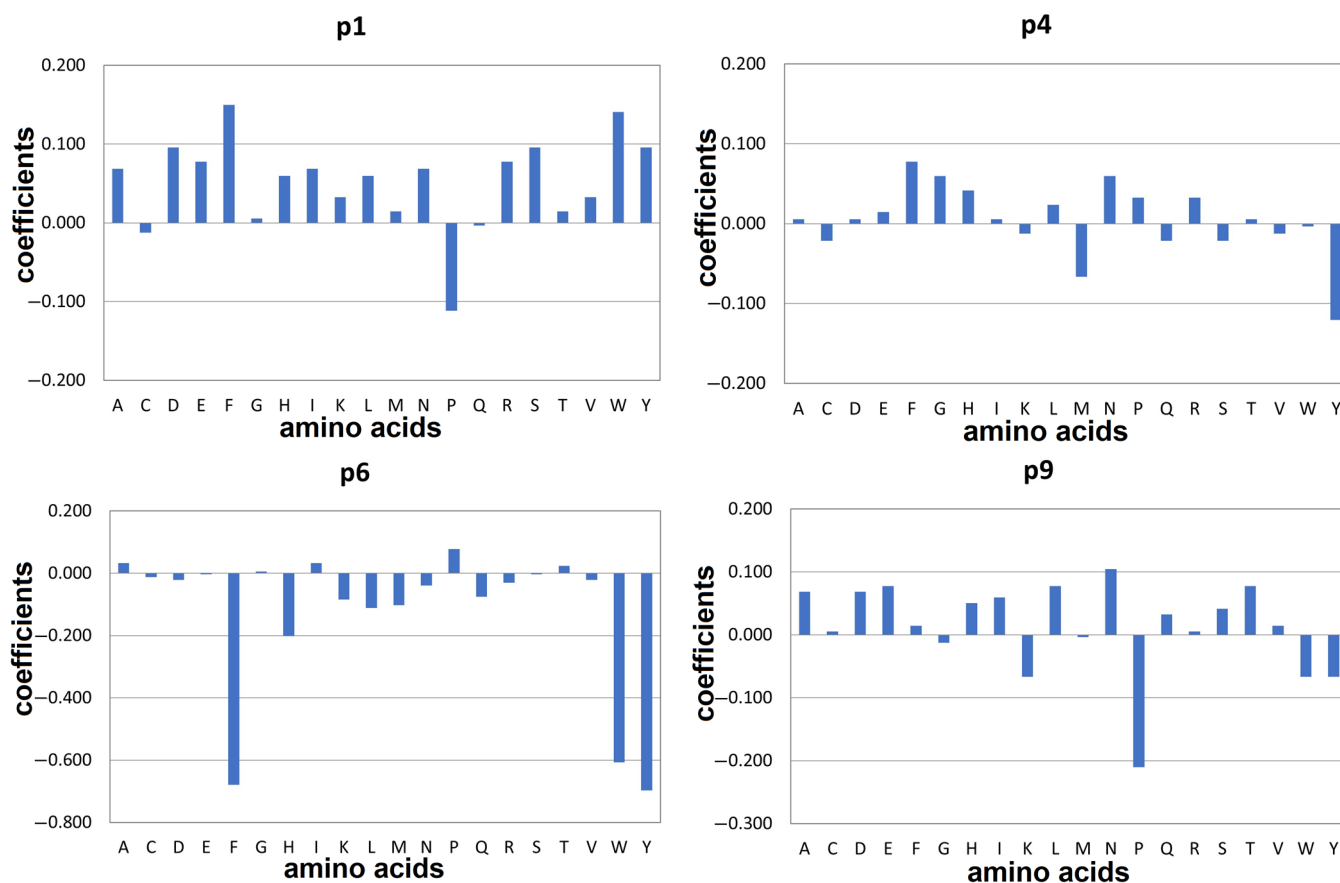


Figure 4. Contributions of the amino acids at anchor positions p1, p4, p6, and p9 according to the QM for α -gliadin peptide library for HLA-DQ8.1. Positive coefficients indicate positive contributions and negative coefficients indicate negative contributions.

3. Discussion

The European Food Safety Authority (EFSA) is responsible for assessing and providing scientific advice on food safety and nutrition-related issues within the European Union. To this end, EFSA plays a crucial role in ensuring the safety of novel food products for individuals with celiac disease [31]. Recently, EFSA has issued guidance on assessing the risk of novel proteins causing celiac disease [32]. The guidance defines a framework that utilizes various methods, such as *in silico*, *in vitro*, and *in vivo* testing, to evaluate the potential of new proteins to trigger an immune response. The novel protein is initially compared to known proteins associated with celiac disease, and if any relevant identities and/or similarities are found, the protein is searched for binders to HLA-DQ alleles associated with celiac disease. If strong peptide binders are predicted, HLA-DQ binding assays, *in vitro* digestibility tests, and tests for T-cell epitopes are conducted to confirm or reject these predictions. This guidance ensures the safety of novel proteins and assists food manufacturers and regulatory agencies in assessing the risk of celiac disease.

Here, following the principles in EFSA's guidance for risk assessment of novel proteins, we describe the development of three docking-based models for the prediction of peptide binding to two HLA-DQ alleles associated with susceptibility to celiac disease. HLA-DQ2.5 is the main such allele available in 80–85% of the patients diagnosed with celiac disease [12]. The peptide binding motif for HLA-DQ2.5 includes preferences for aromatic or hydrophobic aliphatic residues in p1, acidic residues in p4, p6, and/or p7, and hydrophobic or aromatic residues at p9 [24,33]. Additionally, Stepniak et al. [34] found that the T-cell epitopes originating from HLA-DQ2.5 binders preferably carry Pro at p1, p3 and p8 and Glu at p4 and p6. Based on high-throughput mass spectroscopy of HLA ligand elution data, Kosaloglu-Yalcin et al. identified three binding motifs for HLA-DQ2.5: 'motif 1' (Glu

and Asp preferred at p4, p6, and p7), 'motif 2' (Pro at p1, Glu at p7 and p9), and 'no clear motif' (hydrophobic residues at p1, p4, and p6, Glu at p7 and p9) [30]. All these findings indicate that peptides bind to HLA-DQ2.5 in diverse modes, leading to varying preferences for specific residues at anchor positions.

Our studies revealed that amino acid preferences are influenced by the peptide backbone conformation. The superposition of the gliadin and non-gliadin peptides shows that the non-gliadin peptide has a more deeply buried p4 and a more flanking p10 (Figure 5). The presence of a deeply buried p4 explains the preference for small amino acids such as Val, Asp, and Gly at this position (Table S2). On the other hand, the greater freedom at p10 justifies the acceptance of bulky amino acids like Trp, which is not tolerated at the same position in the gliadin peptide due to the rigid poly-Pro backbone, as we demonstrated by MD simulations.

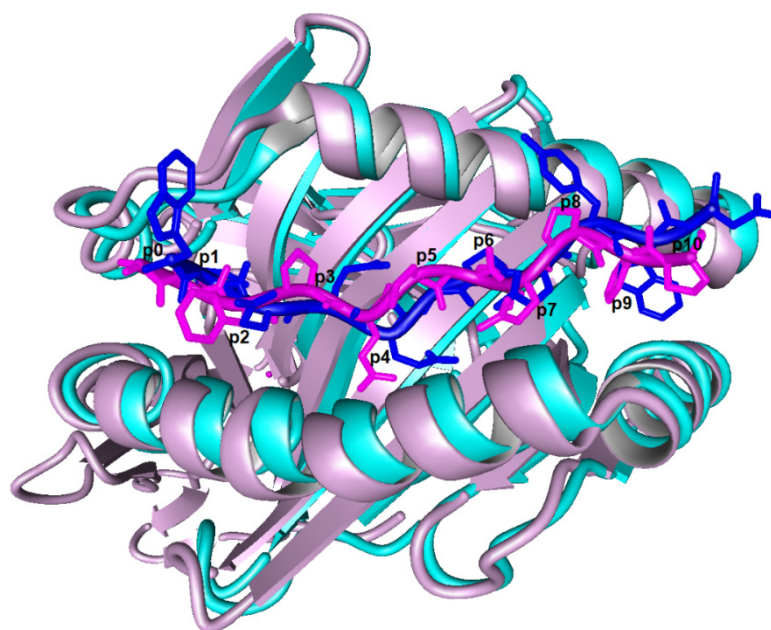


Figure 5. Superposition of α -gliadin peptide QPFPQPELPYP (magenta), HLA-DQ2.5 (mauve), and non-gliadin peptide WIEQEGPEYWD (blue) HLA-DQ2.5 (cyan). The rmsd over peptide backbones is 2.460 Å.

The results of our study align well with the known binding motifs for HLA-DQ2.5 [30], indicating that most of the preferred and non-preferred amino acids are in agreement. Both gliadin and non-gliadin libraries accept aromatic and aliphatic amino acids at p1. However, the presence of Pro at p1 has different effects: it is preferred as part of the poly-Pro backbone in gliadins, while it is detrimental in non-gliadin sequences. Hydrophobic or acidic residues are favoured at p4 in both gliadins and non-gliadins. At p6, significant differences emerge between gliadins and non-gliadins concerning aromatic residues: gliadins prefer them, whereas non-gliadins tend to avoid them. One possible explanation for these differences could be the presence of Pro, this time in the non-gliadin peptide. Substituting the rigid Pro at p6 with bulky aromatic residues may cause unfavourable distortion in the peptide backbone. Hydrophobic amino acids are well-suited for pocket 7. There is a consensus regarding the preferred aromatic amino acids and Glu at p9 between the gliadin and non-gliadin libraries, which aligns with Kosaloglu-Yalcin's observation of 'no clear motif' [30].

The peptide binding motif for HLA-DQ8.1 includes negatively charged amino acids at both p1 and p9, bulky but not basic residues at p4, and small residues at p6 [19,35–38]. The preferences and non-preferences identified in the present study by docking simulations fully agree with the known binding motif. Additionally, our findings reveal that Phe and

Tyr also are well accepted at p1, Pro at p1 and p9, Lys at p4; Phe, Tyr, Arg, and His are unfavoured.

The QMs derived in the present study by docking simulations of combinatorial peptide libraries enable the screening of vast protein libraries, even whole proteomes, significantly reducing time and cost compared to experimental methods. Moreover, they facilitate the exploration of various binding conformations and the prediction of binding affinities. The presented QMs will be implemented in a software tool for predicting peptide binding to alleles associated with celiac disease. In line with the principles in the EFSA's guidance for allergenicity risk assessment of novel proteins, this implementation will enhance the predictive capabilities and aid in the assessment of potential risks.

4. Materials and Methods

4.1. Structures and Combinatorial Libraries

The X-ray structure of HLA-DQ2.5 in complex with the α -gliadin peptide Q₀P₁F₂P₃Q₄P₅E₆L₇P₈Y₉P₁₀ (pdb code: 6mfg [28]) was used in the molecular docking calculations and MD simulations. The non-gliadin endecamer peptide W₀I₁E₂Q₃E₄G₅P₆E₇Y₈W₉D₁₀ [29] was modelled by homology using 6mfg as a starting conformation and geometry optimized by MD simulation for 1000 ns. The crystallographic structure of the complex between HLA-DQ8.1 and α -gliadin peptide S₋₁G₀E₁G₂S₃F₄Q₅P₆S₇Q₈E₉N₁₀P₁₁ (pdb code: 2nna [10]) was used.

The peptides bound in HLA-DQ parent combinatorial libraries constructed by the method of single amino acid substitution, i.e., the residues at each position in the parent peptide were substituted for with the remaining 19 naturally occurring amino acids following the rule 'one peptide–one substitution'. Each library consisted of 210 endecamers (19 amino acids \times 11 positions + 1 parent peptide) and covered the binding core plus the flanking p0 and p10. Positions p–1 and p11 in the peptides for DQ8.1 were kept unchanged.

The external test set used for the validation of the derived QMs for HLA-DQ2.5 consisted of 4249 binders and 4249 non-binders. The DQ2.5 binders were qualitatively measured by mass spectrometry [39]. Only peptides longer than 9 amino acids were included in the set. The set of non-binders was generated by a random combination of non-preferred amino acids at all nine positions from the binding core [30].

The test set for validation of the QMs for HLA-DQ8.1 was constructed similarly to the test set for HLA-DQ2.5. A set of 4339 strong binding nonamers eluted by LC-MS was selected as a positive test set [7]. The same number of non-binders was generated by random combination of non-preferred amino acids in the binding core according to the binding motif for HLA-DQ8.1 [7].

The binding affinities of the tested peptides were calculated as a sum of the amino acids' contributions at each binding core position. The accuracy of predictions was assessed at different cutoffs between binders and non-binders in the range between -0.5 and $+0.5$ with a step of 0.1 , and the cutoff with the highest accuracy was selected.

4.2. Molecular Docking Protocol

Proteins and peptides were pre-processed by MGLTools v. 1.5.7 (Scripps, San Diego, CA, USA). Mutations and visualizations were performed by YASARA (YASARA Biosciences GmbH, Wien, Austria). Each peptide from the combinatorial libraries was docked into the corresponding HLA-DQ protein using the software tool AutoDock Vina v. 1.2.0 [40]. The docking was performed using the AutoDock Vina forcefield, and the resulting binding energy was reported as an affinity in kcal/mol. To optimize the docking protocols, several parameters were considered, including the flexibility of the peptide and protein binding site, exhaustiveness, and energy range. The binding sites were defined to include protein residues within a 6 \AA distance from the bound peptide. AutoDock Vina provided the exhaustiveness parameter, which determines the computational effort used during a docking experiment [41]. The default exhaustiveness value is 8, but in this study, exhaus-

tiveness was tested by increasing it to 30, resulting in more consistent docking results. The energy range parameter defines the range of the best-scored poses [41]. Binding poses with scores outside this range were discarded. In this study, different energy ranges were tested, ranging from 1 to 6 kcal/mol. The best scores obtained for each peptide in the library were mean-normalized, and a quantitative matrix (QM) was constructed. To validate the QMs, external test sets were used, and the best-performing QMs were selected for further analysis.

4.3. MD Simulations Protocol

Initially, the energy of the peptide–HLA complexes was minimized for 10,000 steps using 10 kcal/molÅ² harmonic restraints on solute-heavy atoms and a 10 Å cutoff for the non-bonded van der Waals and electrostatic interactions. Next, the complexes were heated to 300 K in a timestep of 2 fs for 100 ps, density-equilibrated for 100 ps, and equilibrated for 1 ns at constant temperature (300 K) and constant pressure (1 bar). Finally, the production phase was run for 1000 ns (1 μs). Frames were recorded every 1 ns for a total of 1000 per trajectory. The MD simulations in the present study were performed using AMBER 18 [42]. The trajectories were analysed by cpptraj V4.24.0 [43]. Two parameters were calculated: root-mean-square deviation (rmsd), which accounts for changes in atomic coordinates of input frames to a reference frame (frame #1 by default) and rmsfs (root-mean-square fluctuations), which account for atomic positional fluctuations. Three runs were conducted for each complex.

5. Conclusions

The EFSA's role in ensuring food safety, especially for individuals with celiac disease, is evident in their guidance for assessing the risk of novel proteins causing immune reactions linked to celiac disease. Their comprehensive framework, using *in silico*, *in vitro*, and *in vivo* methods, is pivotal in evaluating these risks, requiring comparison to known celiac-associated proteins and subsequent binding assays.

The results found in this study represent a pivotal step in applying the EFSA's risk assessment principles for novel proteins. The derived docking-based models (QMs) for peptide binding to HLA-DQ alleles linked to celiac disease enhance our understanding of peptide–HLA interactions, reveal diverse binding modes, and link specific residue preferences to peptide backbone conformations. Validation of these QMs on known binders and non-binders displayed substantial predictive accuracy (89–94%). QMs demonstrate their efficiency in screening protein libraries, including whole proteomes, significantly reducing time and cost compared to experimental methods. These QMs will be integrated into software for predicting peptide binding, aligning with the EFSA's principles and enhancing risk assessment for proteins triggering celiac disease.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules29010138/s1>, Table S1. Docking-based quantitative matrix for α-gliadin peptide library docked in HLA-DQ2.5; Table S2. Docking-based quantitative matrix for non-gliadin peptide library docked in HLA-DQ2.5; Table S3. Docking-based quantitative matrix for α-gliadin peptide library docked in HLA-DQ8.1. Figure S1a. Accuracy of predictions by the QMs for α-gliadin peptide (blue curve) and non-gliadin peptide (orange curve) at different cutoffs between binders and non-binders to HLA-DQ2.5. Figure S1b. Accuracy of predictions by the QM for HLA-DQ8.1 at different cutoffs between binders and non-binders. Figure S2. Intermolecular interactions between the most positively/negatively contributing peptide residues from α-gliadin and non-gliadin peptides binding into the corresponding pockets of HLA-DQ2.5 and HLA-DQ8.1. The hydrogen bonds are presented as orange discontinued lines; π-π-stacking with red or pale pink for weaker contacts, hydrophobic interactions with green or pale green for weaker interactions. The peptide backbone and sidechains are coloured in cyan. Hydrogens are not presented.

Author Contributions: Conceptualization, M.A., I.D. (Ivan Dimitrov) and I.D. (Irina Doytchinova); methodology, M.A. and I.D. (Irina Doytchinova); validation, I.D. (Ivan Dimitrov); investigation, M.A., I.D. (Ivan Dimitrov) and I.D. (Irina Doytchinova); writing—original draft preparation, I.D. (Irina Doytchinova); writing—review and editing, M.A., I.D. (Ivan Dimitrov), I.D. (Irina Doytchinova), A.F., J.M. and F.K.; visualization, M.A. and I.D. (Irina Doytchinova); supervision, A.F., J.M. and F.K.; project administration, I.D. (Irina Doytchinova); funding acquisition, M.A., I.D. (Ivan Dimitrov) and I.D. (Irina Doytchinova). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Food Safety Authority, grant number OC/EFSA/GMO/2019/02. The in silico calculations were performed in the Centre of Excellence for Informatics and ICT, supported by the Science and Education for Smart Growth Operational Program and co-financed by the European Union through the European Structural and Investment funds (Grant No. BG05M2OP001-1.001-0003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Patronov, A.; Doytchinova, I. T-cell epitope vaccine design by immunoinformatics. *Open Biol.* **2013**, *3*, 120139. [[CrossRef](#)] [[PubMed](#)]
2. Robinson, J.; Barker, D.J.; Georgiou, X.; Cooper, M.A.; Flicek, P.; Marsh, S.G.E. IPD-IMGT/HLA database. *Nucleic Acids Res.* **2020**, *48*, D948–D955. [[PubMed](#)]
3. Raychaudhuri, S.; Sandor, C.; Stahl, E.A.; Freudenberg, J.; Lee, H.S.; Jia, X.; Alfredsson, L.; Padyukov, L.; Klareskog, L.; Worthington, J.; et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.* **2012**, *44*, 291–296. [[CrossRef](#)] [[PubMed](#)]
4. Larid, G.; Pancarte, M.; Offer, G.; Clavel, C.; Martin, M.; Pradel, V.; Auger, I.; Lafforgue, P.; Roudier, J.; Serre, G.; et al. In rheumatoid arthritis patients, HLA-DRB1*04:01 and rheumatoid nodules are associated with ACPA to a particular fibrin epitope. *Front. Immunol.* **2021**, *12*, 692041. [[CrossRef](#)] [[PubMed](#)]
5. Pan, C.F.; Wu, C.J.; Chen, H.H.; Chang, F.M.; Liu, H.F.; Chu, C.C.; Lin, M.; Lee, Y.J. Molecular analysis of HLA-DRB1 allelic associations with systemic lupus erythematosus and lupus nephritis in Taiwan. *Lupus* **2009**, *18*, 698–704. [[CrossRef](#)] [[PubMed](#)]
6. Morris, D.L.; Taylor, K.E.; Fernando, M.M.A.; Nititham, J.; Alarcón-Riquelme, M.E.; Barcellos, L.F.; Behrens, T.W.; Cotsapas, C.; Gaffney, P.M.; Graham, R.R.; et al. Unraveling multiple MHC gene associations with systemic lupus erythematosus: Model choice indicates a role for HLA alleles and non-HLA genes in Europeans. *Am. J. Hum. Genet.* **2012**, *91*, 778–793. [[CrossRef](#)] [[PubMed](#)]
7. Tran, M.T.; Faridi, P.; Lim, J.J.; Ting, Y.T.; Onwukwe, G.; Bhattacharjee, P.; Jones, C.M.; Tresoldi, E.; Cameron, F.J.; La Gruta, N.L.; et al. T cell receptor recognition of hybrid insulin peptides bound to HLA-DQ8. *Nat. Commun.* **2021**, *12*, 5110. [[CrossRef](#)]
8. Zhao, L.P.; Papadopoulos, G.K.; Moustakas, A.K.; Bondinas, G.P.; Carlsson, A.; Larsson, H.E.; Ludvigsson, J.; Marcus, C.; Persson, M.; Samuelsson, U.; et al. Nine residues in HLA-DQ molecules determine with susceptibility and resistance to type 1 diabetes among young children in Sweden. *Sci. Rep.* **2021**, *11*, 8821. [[CrossRef](#)]
9. Fontenot, A.P.; Kotzin, B.L. Chronic beryllium disease: Immune-mediated destruction with implications for organ-specific autoimmunity. *Tissue Antigens* **2003**, *62*, 449–458. [[CrossRef](#)]
10. Henderson, K.N.; Tye-Din, J.A.; Reid, H.H.; Chen, Z.; Borg, N.A.; Beissbarth, T.; Tatham, A.; Mannering, S.I.; Purcell, A.W.; Dudek, N.L.; et al. A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. *Immunity* **2007**, *27*, 23–34. [[CrossRef](#)]
11. Cecilio, L.A.; Bonatto, M.W. The prevalence of HLA DQ2 and DQ8 in patients with celiac disease, in family and in general population. *Arq. Bras. De Cir. Dig.* **2015**, *28*, 183–185. [[CrossRef](#)] [[PubMed](#)]
12. Ludvigsson, J.F.; Leffler, D.A.; Bai, J.C.; Biagi, F.; Fasano, A.; Green, P.H.; Hadjivassiliou, M.; Kaukinen, K.; Kelly, C.P.; Leonard, J.N.; et al. The Oslo definitions for coeliac disease and related terms. *Gut* **2013**, *62*, 43–52. [[CrossRef](#)] [[PubMed](#)]
13. Kagnoff, M.F. Celiac Disease: Pathogenesis of a Model Immunogenetic Disease. *J. Clin. Investig.* **2007**, *117*, 41–49. [[CrossRef](#)] [[PubMed](#)]
14. Elli, L.; Branchi, F.; Tomba, C.; Villalta, D.; Norsa, L.; Ferretti, F.; Roncoroni, L.; Bardella, M.T. Diagnosis of gluten related disorders: Celiac disease, wheat allergy and non-celiac gluten sensitivity. *World J. Gastroenterol.* **2015**, *21*, 10331–10341. [[CrossRef](#)] [[PubMed](#)]
15. Lebowitz, B.; Ludvigsson, J.F.; Green, P.H. Celiac disease and non-celiac gluten sensitivity. *BMJ* **2015**, *351*, h4347. [[CrossRef](#)] [[PubMed](#)]
16. Ghazanfar, H.; Javed, N.; Lee, S.; Shaban, M.; Cordero, D.; Acherjee, T.; Hasan, K.Z.; Jyala, A.; Kandhi, S.; Hussain, A.N.; et al. Novel Therapies for Celiac Disease: A Clinical Review Article. *Cureus* **2023**, *15*, e39004. [[CrossRef](#)] [[PubMed](#)]

17. Gonzalez-Galarza, F.F.; McCabe, A.; Santos, E.J.; Jones, J.; Takeshita, L.Y.; Ortega-Rivera, N.D.; Del Cid-Pavon, G.M.; Ramsbottom, K.; Ghattaoraya, G.S.; Alfirevic, A.; et al. Allele frequency net database (AFND) 2020 update: Gold-standard data classification, open access genotype data and new query tools. *Nucleic Acid Res.* **2020**, *48*, D783–D788. [[CrossRef](#)]
18. Liu, E.; Rewers, M.; Eisenbarth, G.S. Genetic Testing: Who Should Do the Testing and What Is the Role of Genetic Testing in the Setting of Celiac Disease? *Gastroenterology* **2005**, *128*, S33–S37. [[CrossRef](#)]
19. Moustakas, A.K.; van de Wal, Y.; Routsias, J.; Kooy, Y.M.C.; van Veelen, P.; Drijfhout, J.W.; Koning, F.; Papadopoulos, G.K. Structure of Celiac Disease-Associated HLA-DQ8 and Non-Associated HLA-DQ9 Alleles in Complex with Two Disease-Specific Epitopes. *Int. Immunol.* **2000**, *12*, 1157–1166. [[CrossRef](#)]
20. van de Wal, Y.; Kooy, Y.M.C.; Drijfhout, J.W.; Amons, R.; Papadopoulos, G.K.; Koning, F. Unique Peptide Binding Characteristics of the Disease-Associated DQ(α 1*0501, β 1*0201) vs the Non-Disease-Associated DQ(α 1*0201, β 1*0202) Molecule. *Immunogenetics* **1997**, *46*, 484–492. [[CrossRef](#)]
21. Kumar, R.; Eastwood, A.L.; Brown, M.L.; Laurie, G.W. Human Genome Search in Celiac Disease: Mutated Gliadin T-Cell-like Epitope in Two Human Proteins Promotes T-Cell Activation. *J. Mol. Biol.* **2002**, *319*, 593–602. [[CrossRef](#)] [[PubMed](#)]
22. Costantini, S.; Rossi, M.; Colonna, G.; Facchiano, A.M. Modelling of HLA-DQ2 and Its Interaction with Gluten Peptides to Explain Molecular Recognition in Celiac Disease. *J. Mol. Graph. Model.* **2005**, *23*, 419–431. [[CrossRef](#)] [[PubMed](#)]
23. Gianfrani, C.; Siciliano, R.A.; Facchiano, A.M.; Camarca, A.; Mazzeo, M.F.; Costantini, S.; Salvati, V.M.; Maurano, F.; Mazzarella, G.; Iaquinto, G.; et al. Transamidation of Wheat Flour Inhibits the Response to Gliadin of Intestinal T Cells in Celiac Disease. *Gastroenterology* **2007**, *133*, 780–789. [[CrossRef](#)] [[PubMed](#)]
24. Vartdal, F.; Johansen, B.H.; Friede, T.; Thorpe, C.J.; Stevanovic, S.; Eriksen, J.E.; Sletten, K.; Thorsby, E.; Rammensee, H.G.; Sollid, L.M. The peptide binding motif of the disease associated HLA-DQ (α 1*0501, β 1*0201) molecule. *Eur. J. Immunol.* **1996**, *26*, 2764–2772. [[CrossRef](#)] [[PubMed](#)]
25. Barre, A.; Simplicien, M.; Cassan, G.; Benoist, H.; Rougé, P. Docking of Peptide Candidates to HLA-DQ2 and HLA-DQ8 Basket as a Tool for Predicting Potential Immunotoxic Peptides toward Celiac Diseased People. *Rev. Fr. Allergol.* **2018**, *58*, 482–488. [[CrossRef](#)]
26. Ting, Y.T.; Dahal-Koirala, S.; Kim, H.S.K.; Qiao, S.-W.; Neumann, R.S.; Lundin, K.E.A.; Petersen, J.; Reid, H.H.; Sollid, L.M.; Rossjohn, J. A Molecular Basis for the T Cell Response in HLA-DQ2.2 Mediated Celiac Disease. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 3063–3073. [[CrossRef](#)] [[PubMed](#)]
27. Atanasova, M.; Patronov, A.; Dimitrov, I.; Flower, D.R.; Doytchinova, I. EpiDOCK: A Molecular Docking-Based Tool for MHC Class II Binding Prediction. *Protein Eng. Des. Sel.* **2013**, *26*, 631–634. [[CrossRef](#)]
28. Dahal-Koirala, S.; Ciacchi, L.; Petersen, J.; Risnes, L.F.; Neumann, R.S.; Christophersen, A.; Lundin, K.; Reid, H.H.; Qiao, S.W.; Rossjohn, J.; et al. Discriminative T-cell receptor recognition of highly homologous HLA-DQ2-bound gluten epitopes. *J. Biol. Chem.* **2019**, *294*, 941–952. [[CrossRef](#)]
29. van de Wal, Y.; Kooy, Y.M.; Drijfhout, J.W.; Amons, R.; Koning, F. Peptide binding characteristics of the coeliac disease-associated DQ (α 1*0501, β 1*0201) molecule. *Immunogenetics* **1996**, *44*, 246–253. [[CrossRef](#)]
30. Koşaloğlu-Yalçın, Z.; Sidney, J.; Chronister, W.; Peters, B.; Sette, A. Comparison of HLA ligand elution data and binding predictions reveals varying prediction performance for the multiple motifs recognized by HLA-DQ2.5. *Immunology* **2021**, *162*, 235–247. [[CrossRef](#)]
31. Commission Implementing Regulation (EU) No 503/2013 of 3 April 2013 on applications for authorisation of genetically modified food and feed in accordance with Regulation (EC) No 1829/2003 of the European Parliament and of the Council and amending Commission Regulations (EC) No 641/2004 and (EC) No 1981/2006. OJ L157, 8.6.2013. pp. 1–48. Available online: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:157:0001:0048:EN:PDF> (accessed on 28 November 2023).
32. EFSA Panel on Genetically Modified Organisms (GMO); Naegeli, H.; Birch, A.N.; Casacuberta, J.; De Schrijver, A.; Gralak, M.A.; Guerche, P.; Jones, H.; Manachini, B.; Messéan, A.; et al. Guidance on allergenicity assessment of genetically modified plants. *EFSA J.* **2017**, *15*, e04862.
33. Sidney, J.; Steen, A.; Moore, C.; Ngo, S.; Chung, J.; Peters, B.; Sette, A. Divergent motifs but overlapping binding repertoires of six HLA-DQ molecules frequently expressed in the worldwide human population. *J. Immunol.* **2010**, *185*, 4189–4198. [[CrossRef](#)]
34. Stepniak, D.; Wiesner, M.; de Ru, A.H.; Moustakas, A.K.; Drijfhout, J.W.; Papadopoulos, G.K.; van Veelen, P.A.; Koning, F. Large-scale characterization of natural ligands explains the unique gluten-binding properties of HLA-DQ2. *J. Immunol.* **2008**, *180*, 3268–3278. [[CrossRef](#)] [[PubMed](#)]
35. Kwok, W.W.; Nepom, G.T.; Raymond, F.C. HLA-DQ polymorphisms are highly selective for peptide binding interactions. *J. Immunol.* **1995**, *155*, 2468–2476. [[CrossRef](#)] [[PubMed](#)]
36. Kwok, W.W.; Domeier, M.E.; Johnson, M.L.; Nepom, G.T.; Koelle, D.M. HLA-DQB1 codon 57 is critical for peptide binding and recognition. *J. Exp. Med.* **1996**, *183*, 1253–1258. [[CrossRef](#)] [[PubMed](#)]
37. Kwok, W.W.; Domeier, M.E.; Raymond, F.C.; Byers, P.; Nepom, G.T. Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. *J. Immunol.* **1996**, *156*, 2171–2177. [[CrossRef](#)] [[PubMed](#)]
38. Andreatta, M.; Nielsen, M. Characterizing the binding motifs of 11 common human HLA-DP and HLA-DQ molecules using NNAlign. *Immunology* **2012**, *136*, 306–311. [[CrossRef](#)]

39. Bergseng, E.; Dørum, S.; Arntzen, M.Ø.; Nielsen, M.; Nygård, S.; Buus, S.; de Souza, G.A.; Sollid, L.M. Different binding motifs of the celiac disease-associated HLA molecules DQ2.5, DQ2.2, and DQ7.5 revealed by relative quantitative proteomics of endogenous peptide repertoires. *Immunogenetics* **2015**, *67*, 73–84. [[CrossRef](#)]
40. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
41. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, *61*, 3891–3898. [[CrossRef](#)]
42. Case, D.A.; Cheatham, T.E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688. [[CrossRef](#)] [[PubMed](#)]
43. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.