Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 23 (2005) 419–431
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Modelling of HLA-DQ2 and its interaction with gluten peptides to explain molecular recognition in celiac disease

Susan Costantini[a,b], Mauro Rossi[a], Giovanni Colonna[b], Angelo M. Facchiano[a,b,*]

[a]*Institute of Food Science, CNR, via Roma 52 A/C, 83100 Avellino, Italy*
[b]*CRISCEB, Research Center of Computational and Biotechnological Sciences, Second University of Naples, via Costantinopoli 16, 80138 Naples, Italy*

Received 26 July 2004; received in revised form 8 November 2004; accepted 16 December 2004
Available online 18 January 2005

Abstract

Celiac disease (CD) is sustained by abnormal intestinal mucosal T-cell response to gluten and it is strongly associated with HLA class II molecules encoded by DQA1 *0501/DQB1 *02 (DQ2) or DQA1 *03/DQB1 *0302 (DQ8). The in vitro stimulatory activity of gliadin increases after treatment with tissue transglutaminase (tTG) which catalyses the deamidation of specific residues of glutamine to glutamate that can serve as anchors for binding to DQ2 as well as to DQ8 molecules. We modelled the three-dimensional structure of the DQ2 dimer protein, the most frequent in celiac patients, by using a homology modelling strategy, and deposited the model in the Protein Data Bank (PDB). Then, we simulated the interactions of DQ2 with different gluten peptides and the deamidation of specific peptide glutamines in the known p4, p6, p7 and p9 anchor positions, as well as in p1 and p5 positions, and other substitutions for which experimental effects on binding are available by previous experimental studies. By evaluating the energy of interaction and the H-bond interactions, we were able to distinguish what substitutions improve the interaction peptide–DQ2, in agreement with previously published experimental data. By analysing the peptide–DQ2 complex at the atom level, we observed that these glutamate side chains can interact with specific positively charged amino acids of DQ2, absent in other HLA alleles not related to celiac disease. The simulation was also extended to other peptides, related to celiac disease but for which no experimental data exists about the effects of glutamine deamidation. Our results give an interpretation at the molecular level of previously reported binding experimental data and open a new window to gain further insights about peptide recognition in celiac disease.

Keywords: Celiac disease; HLA-DQ2; Gluten peptides; Protein–peptide interaction; Molecular simulations; MHC class II

1. Introduction

Celiac disease (CD), or celiac sprue, or gluten-sensitive enteropathy, is a multifactorial disorder influenced by both environmental and genetic factors. It is characterized by malabsorption resulting from inflammatory injury to the mucosa of the small intestine after the ingestion of wheat gluten or related rye and barley proteins. It often starts shortly after the first introduction of wheat into the diet, and symptoms include diarrhoea, malabsorption, and failure to thrive, which is due to an inefficient uptake of nutrients by a flattening intestinal epithelium. At present the only effective treatment for the disease is the removal of gluten from the diet, but the reintroduction of gluten in the patient's diet invariably leads to the reappearance of the symptoms [1–3]. Although the molecular basis of CD is still unclear, the molecular mechanism is considered to involve the binding of gluten peptides to HLA molecules and then the specific recognition by T-cells [2]. Different gluten-derived peptides have been identified as able to be recognized by T-cell clones isolated from biopsies of CD patients [4–6]. HLA-DQ2 (DQA1 *0501/B1 *0201) is found in the great majority of CD patients, while DQ8 (DQA1 *0301/B1 *0302) is found in most of the remaining patients [1,2]. The binding of gluten peptides to DQ2 and DQ8 molecules has been experimentally observed and it may be improved by the presence of

Abbreviations: 3D, three-dimensional; CD, celiac disease; HLA, human leukocyte antigen; PDB, Protein Data Bank; tTG, tissue transglutaminase

* Corresponding author. Tel.: +39 0825 299625; fax: +39 0825 299813.
E-mail address: angelo.facchiano@isa.cnr.it (A.M. Facchiano).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.12.002

negatively charged amino acids at known positions in the peptide [1]. Gluten proteins are very rich of glutamine and proline and contain very few negatively charged amino acids, but glutamines may be deamidated to the negatively charged glutamate by the tissue transglutaminase (tTG) enzyme, with the consequence of improving both the binding to DQ2 and the response of T-cell clones [7–9].

Different studies have been carried out in order to simulate the interaction of gluten peptides with DQ2. Some simulations were started from the knowledge of the experimental model of the DR1 dimer and based on models of the DQ2 molecule obtained by simple modelling procedures as the amino acids substitutions [10,11]. In two other works, a probably more complete but not described homology modelling procedure was applied starting from the DR1 [12] or DQ8 dimer [13]. The assessment of more efficient modelling procedures as well as the availability of the experimental structure of the DQ8 dimer, which is more similar to DQ2 and represents a better template model than DR1, makes it possible to improve the methods previously applied and to perform more accurate simulations of the interaction of gluten peptides with DQ2, as well as an accurate analysis of the effect of deamidation and other modifications of the peptides.

In our study, we used bioinformatics and biocomputing approaches and tools for predicting the 3D structure of DQ2 molecule, validate it and deposit the model in the PDB [14], in order to make the model available to the scientific community. Then, we simulated the complex between DQ2 and different gluten-derived peptides and their modified forms, to investigate the molecular details of this interaction and the effects of modification of specific peptide amino acids, in order to propose an interpretation at the molecular level of published experimental results concerning the binding of gluten peptides to DQ2 and the effects of deamidation. We also applied the same strategy to other gluten peptides for which a relation to celiac disease has been established, but no experimental evidence exists about deamidation of glutamines. Finally, to find further information about the specificity of DQ2 recognition, we analysed the presence of functional groups in the peptide-binding site, and evidenced the relevance of DQ2 structural peculiarities in comparison to other DQ alleles not related to celiac disease. We also discuss the results of our computational work in comparison to the very recently published article describing the experimental structure of the complex between DQ2 and an immunogenic epitope from gluten [15].

2. Methodology

2.1. Protein modelling

Protein sequences of DQ2 chains used for protein modelling were selected from SwissProt database (DQA1 *0501 chain: accession number P01909; DQB1 *0201 chain: accession number P01918). The 3D model of DQ2 was created following the homology modelling procedure [16–19] already used with success by our group [20–24]. In brief, searches for sequence similarity within databases were performed with the BLAST program [25]. The searches evidenced that a very high identity (91%) exists for both DQ2 chains with the corresponding chains in DQ8, for which the 3D structure is available [26], so that the homology modelling strategy can be applied with success by using the DQ8 experimental model as template (PDB code: 1JK8). Due to the lack of terminal segments in both chains of DQ8 model, it was possible to build models of the 4–183 region for the alpha chain and the 3–192 region for the beta chain of DQ2. The programs MODELLER [27] and Quanta (Accelrys, San Diego, CA) were used to build 10 models for each protein chain under prediction. To select the best model among them, their stereo chemical quality was evaluated with the program PROCHECK [28], in particular by comparing the Ramachandran plots and the related statistical analysis. Search for structural classification was performed in CATH [29–30] and SCOP [31,32] databases. The models were analyzed with different programs to analyze structural characteristics, i.e. DSSP [33] to assign secondary structure, NACCESS [34] to evaluate solvent accessibility.

Finally, the quaternary structure of the DQ2 dimer was assembled by superimposing the two chains to the DQ8 chains. The CVFF force field within the Discover module of InsightII (Accelrys, Inc.) was used to assign potentials and charges, and a mild energy minimization was applied in vacuo (dielectric constant = 1.00), in order to optimise the interaction between the two protein chains and avoid sterical clashes, by performing 500 steps under conjugate gradient algorithm.

2.2. Simulation of DQ2–peptide complexes

The following gluten peptides, known to stimulate in vitro T-cells isolated by DQ2 celiac patients, were used for simulations: alpha-I (fragment 60–68 from alpha 9 gliadin) [6], alpha-II (fragment 62–70 from alpha 2 gliadin) [6], alpha-III (fragment 67–75 from alpha 2 gliadin) [37], glia-a20 (fragment 74–82 from alpha 9 gliadin) [38,39], glia-g2 (fragment 72–80 from gamma 5 gliadin) [38,39], gamma-I (fragment 115–123 from gamma 5 gliadin) [4], gamma-II (fragment 228–236 from gamma 5 gliadin) [37], gamma-III (fragment 66–78 from gamma 5 gliadin) [37], gamma-IV (fragment 102–113 from gamma 5 gliadin) [37], glt-156 (fragment 45–54 from LMW glutenin) [38], glt-17 (fragment 105–113 from LMW glutenin) [38,39]. In addition, we simulated the interaction with many modified forms of these peptides, in order to investigate the effects of deamidation of specific glutamine residues as well as of other substitutions.

The assembly of DQ2–peptide complexes was based on procedures described in previous works [10–13]. In more detail, on the basis of the DQ8– (B9-23) insulin peptide

complex (PDB code: 1JK8), we identified the region of DQ2 involved in the peptide binding and the anchor positions [1]. Each DQ2–peptide complex was created by superimposition of DQ2 to the DQ8 coordinates in the DQ8–peptide complex, and subsequent substitution of the insulin peptide residues with the gluten peptide residues. Each DQ2–peptide complex was then optimised with the software InsightII by using 500 steps of energy minimization under conjugate gradient algorithm, according to the procedure utilized in previous similar works [5,10] and similarly to the described procedure for the DQ2 assembly. This mild minimization was applied to optimise side chain conformations and avoid sterical clashes. After minimization, the energy of interaction between the peptide and the DQ2 molecule was evaluated by using the Energy/Intermolecular tool in the Docking module of InsightII. The HBplus package [35] was used to evaluate the putative formation of H-bonds, which identifies H-bonds within a distance of 2.5 Å and a minimum angle of 90°. Starting from its results, we selected only the stronger H-bonds by restricting the distance to 2.2 Å, according to a recent analysis of geometrical characteristics of H-bonds in protein–ligand complexes [36].

3. Results

3.1. Homology modelling of DQ2

The sequences of both DQ2 chains have been analyzed by computer programs in order to find similar sequences in databases and perform structural predictions to obtain a theoretical 3D model of the protein. The sequences of both DQ8 chains were found to be the most similar to the corresponding DQ2 chains, in terms of sequence similarity, having 91% of sequence identity for both alpha and beta chains (see Fig. 1). In these conditions, a homology modelling strategy can be applied with very good results [16–19], because the alignment between the new sequence and the reference model can be obtained with few or no gaps, in this case just one gap for the alignment of alpha chains and no gap for beta chains. Therefore, the models of DQ2 alpha and beta chains were created by using the homology modelling strategy on the template of the corresponding DQ8 chains (PDB code: 1JK8). The assembly of DQ2 dimer was obtained by superimposing the two best-modelled chains to the corresponding DQ8 chains, in order to obtain the same relative orientation of the two subunits (see under

```
Alpha chains

     4
     |
DQ2  VADHVASYGV  NLYQSYGPSG  QYTHEFDGDE  QFYVDLGRKE  TVWCLPVLRQ
DQ8  VADHVASYGV  NLYQSYGPSG  QYSHEFDGDE  EFYVDLERKE  TVWQLPLFRR

DQ2  F-RFDPQFALT NIAVLKHNLN  SLIKRSNSTA  ATNEVPEVTV  FSKSPVTLGQ
DQ8  FRRFDPQFALT NIAVLKHNLN  IVIKRSNSTA  ATNEVPEVTV  FSKSPVTLGQ

DQ2  PNILICLVDN  IFPPVVNITW  LSNGHSVTEG  VSETSFLSKS  DHSFFKISYL
DQ8  PNTLICLVDN  IFPPVVNITW  LSNGHSVTEG  VSETSFLSKS  DHSFFKISYL

                                     183
                                      |
DQ2  TLLPSAEESY  DCKVEHWGLD  KPLLKHWEPE
DQ8  TFLPSDDEIY  DCKVEHWGLD  EPLLKHWEPE


Beta chains

     3
     |
DQ2  SPEDFVYQFK  GMCYFTNGTE  RVRLVSRSIY  NREEIVRFDS  DVGEFRAVTL
DQ8  SPEDFVYQFK  GMCYFTNGTE  RVRLVTRYIY  NREEYARFDS  DVGVYRAVTP

DQ2  LGLPAAEYWN  SQKDILERKR  AAVDRVCRHN  YQLELRTTLQ  RRVEPTVTIS
DQ8  LGPPAAEYWN  SQKEVLERTR  AELDTVCRHN  YQLELRTTLQ  RRVEPTVTIS

DQ2  PSRTEALNHH  NLLVCSVTDF  YPAQIKVRWF  RNDQEETAGV  VSTPLIRNGD
DQ8  PSRTEALNHH  NLLVCSVTDF  YPAQIKVRWF  RNDQEETTGV  VSTPLIRNGD

                                              192
                                               |
DQ2  WTFQILVMLE  MTPQRGDVYT  CHVEHPSLQS  PITVEWRAQS
DQ8  WTFQILVMLE  MTPQRGDVYT  CHVEHPSLQN  PIIVEWRAQS
```

Fig. 1. Alignment of alpha and beta chains of DQ2 and DQ8. In bold are represented amino acids located into the peptide-binding site, underlines evidence the differences between DQ2 and DQ8.

Section 2). The final molecular model was submitted to the Protein Data Bank [14] and accepted (PDB code: 1NBN) having passed all the conformational parameters checks. Fig. 2A shows the DQ2 structure with evidenced the secondary structure elements. The two chains have a classical organization in two distinct domains. Both N-terminal domains consist of a helical region and a beta sheet of four strands in antiparallel orientation. Both C-terminal domains are characterized by an immunoglobulin-like beta-sandwich made of two antiparallel sheets, each consisting of three main strands and few shorter strands, organized in Greek-keys motifs. By similarity to the DQ8 structure, we assume that the binding site of the peptide ligand consists of a groove at the interface of the N-terminal domains of the two chains, having the two sheets at the bottom of the cavity and the two helical regions as edges. The alpha chain of DQ2 is characterized by two cysteines at the distance suitable to form an S–S bond between the two sheets of the C-terminal domain sandwich architecture, similarly to the DQ8 template model. The third cysteine, not involved in disulphide bonds in our model, is absent in DQ8 and many other alleles. The four cysteines present in beta chain are located at the same position as in DQ8 and may form an S–S bond in the C-terminal domain between the two sheets of the sandwich architecture, and another S–S bond in the N-terminal domain.

The amino acids in the peptide-binding site have been identified and compared between DQ2 and DQ8 (see Fig. 1). From each chain, about 20 amino acids contribute to create the site, and few differences exist between the two HLA molecules. The most evident difference concerns the number and the type of charged side chains. By considering, from DQ2 to DQ8, the substitutions Gln α34 -> Glu, Lys β71 -> Thr, and Arg β77 -> Thr, and the insertion of Arg α55 in DQ8, the peptide-binding site of DQ2 has one more positive charge and one less negative charge, in comparison to DQ8. This difference may be relevant for the ability to bind peptides with negatively charged amino acids (see below).

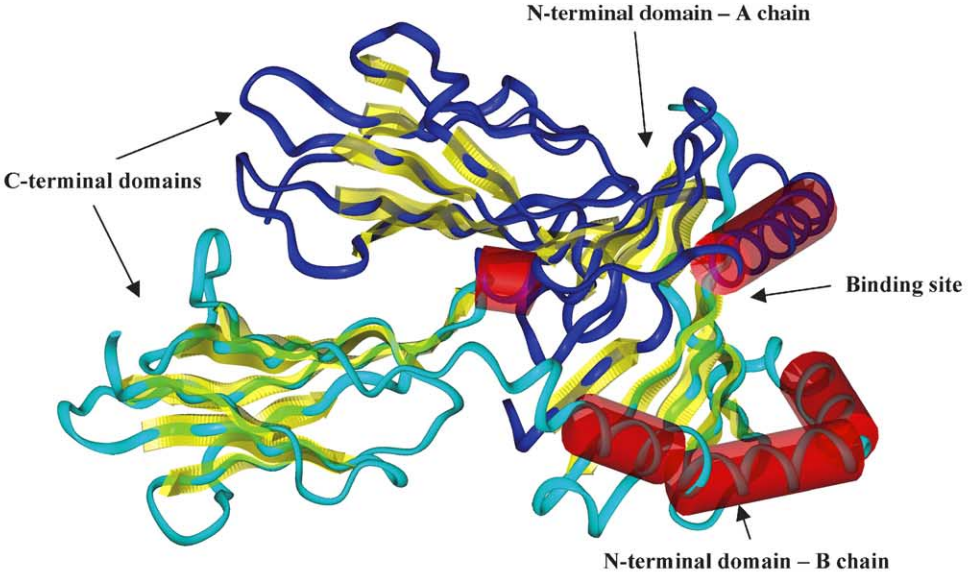
It is very recent the publication of an article describing the experimental structure of DQ2 complexed with an immunogenic epitope from gluten [15]. The availability of an experimental model allows us to verify the quality of our theoretical model of DQ2. The secondary structures assigned to the theoretical and experimental models resulted very similar and very well overlapped (Fig. 2B). By applying structural superposition (see Fig. 2C) and RMSD evaluations, our model appears very similar to the experimental one. The atom-by-atom comparison was not simple because the experimental model is incomplete. In fact, chain beta is interrupted and the segment 105–112 is lacking. We calculated the RMSD values for the structurally conserved regions in the whole dimer (1.08 Å), as well as for each single chain or fragment and for the peptide, obtaining values from 0.5 to 0.7 Å, thus proving the good similarity between our model and the experimental one now available. In Fig. 2D, it is shown the superposition of amino acids recognized by the X-ray study as involved in the interaction with the peptide. A good overlap is evident for all backbone components of amino acids, while differences at side chain level may be considered not relevant (see under Section 4).

Therefore, we consider that the use of our theoretical model for the studies described below is fully valid, because the completeness of the chains makes our model more suitable for energy computation and molecular simulations.

3.2. Simulation of DQ2–peptide interaction and evaluation of effects of side chain modifications

We created structural models of complexes between DQ2 and each peptide. The anchor positions known by previous studies and reported in literature have been considered in order to find the better position of the peptide into the binding site. Different DQ2-restricted gluten peptides are reported in literature as related to celiac disease. For most of them, deamidation of specific glutamine was suggested as a reaction which increases the affinity for DQ2 [7]. In our simulation studies, we used alpha-I, -II, and -III, glia-a20, glia-g2, gamma-I, -II, -III and -IV gliadin-derived peptides, glt-156 and glt-17 glutenin-derived peptides. The sequences of these peptides were aligned to the insulin peptide complexed with the reference model of DQ8 (data not shown), on the basis of literature information concerning the location of specific residues in anchor sites [1,4,6,23,37,39]. Moreover, we simulated the deamidation of different glutamines and other substitutions, to verify the agreement of our models to previously reported experimental results. After energy minimization of each DQ2–peptide complex, we evaluated the putative H-bonds and the energy of interaction, given by the sum of Van der Waals and electrostatic contributions, in order to explain experimental affinity modifications after deamidation or other substitutions. In Fig. 3, we show the differences of energies of interaction between the natural peptide and each mutated peptide. Substitutions have been simulated to verify the agreement of such energy computations to the published experimental results, as well as to investigate the effects of mutations in other anchor positions. In the whole set of data, the difference of energy of interaction are mainly due to the change of electrostatic energy, while Van der Waals contributions appear to be quite unaffected by amino acid modifications (data not shown). This is in agreement with expectable behaviour, in particular when deamidation of glutamine is simulated, by considering the subtle difference existing between glutamine and glutamate side chain. More consistent differences are observed for the electrostatic energy contribution. Substitution of glutamine or other neutral amino acids by the negatively charged glutamate improves in all cases the electrostatic energy of interaction, and multiple substitutions give effects much more evident. On the contrary, the substitution of glutamine with the positively charged lysine gives a loss of electrostatic energy contribution to the total interaction energy. Substitutions

(A)



(B)

```
A Chain
                  10        20        30        40        50        60        70
                   |         |         |         |         |         |         |
SEQUENCE     VADHVASYGVNLYQSYGPSGQYTHEFDGDEQFYVDLGRKETVWCLPVLRQFRFDPQFALTNIAVLKH
MODEL          S EEEEEEEEEETTTEEEEEEEETTEEEEEEETTTTEEEESSHHHHTT   HHHHHHHHHHHHH
1S9V           SEEEEEEEEEEETTTTEEEEEEEETTEEEEEEETTTTEEEESSGGGGGS   THHHHHHHHHHHH

                  80        90        100       110       120       130       140
                   |         |         |         |         |         |         |
SEQUENCE NLNSLIKRSNSTAATNEVPEVTVFSKSPVTLGQPNILICLVDNIFPPVVNITWLSNGHSVTEGVSETSFL
MODEL    HHHHHHHHTTS        EEEEEESS   TTS EEEEEEEEEE SS  EEEEEETTEEE TTEEE   E
1S9V     HHHHHHHHTTS        EEEEEESS   TTS EEEEEEEEEE SS  EEEEEETTEE  SSEEE   E

                  150       160       170       180
                   |         |         |         |
SEQUENCE SKSDHSFFKISYLTLLPSAEESYDCKVEHWGLDKPLLKHW EPE
MODEL    E TTS EEEEEEEEE   SS  EEEEEE TTSSS EEEEE
1S9V     E TTS EEEEEEEEE   SS  EEEEEE TT SS EEEE ---

B Chain
                  10        20        30        40        50        60        70
                   |         |         |         |         |         |         |
SEQUENCE   SPEDFVYQFKGMCYFTNGTERVRLVSRSIYNREEIVRFDSDVGEFRAVTLLGLPAAEYWNSQKDILER
MODEL            EEEEEEEEEETTTTEEEEEEEEEETTEEEEEEETTT SEEESSTTSHHHHHHHHT HHHHHH
1S9V           EEEEEEEEEEETTTTEEEEEEEEEETTEEEEEEETTT SEEESSGGGHHHHHHHHT HHHHHH

                  80        90        100       110       120       130       140
                   |         |         |         |         |         |         |
SEQUENCE KRAAVDRVCRHNYQLELRTTLQRRVEPTVTISPS RTEALNHH NLLVCSVTDFYPAQIKVRWFRNDQEETA
MODEL    HHHHHHHTHHHHHHHHHHTTTT  B  EEEEEE  SS  SSS EEEEEEEEEBSS  EEEEEETTEE SS
1S9V     HHHHHHHTHHHHHHHHHHTGGG  B  EEEEEE -------- EEEEEEEEEBSS  EEEEEETTEEETT

                  150       160       170       180
                   |         |         |         |
SEQUENCE GVVSTPLIRNGDWTFQILVMLEMTPQRGDVYTCHVEHPSLQSPITVEWRA QS
MODEL     EEE   EE SSS EEEEEEEE    TT EEEEEEE TT SS EEEEEES
1S9V      EEE   EE SSS EEEEEEEE    TT EEEEEEE TT SS EEEEEE --
```

Fig. 2. Model of DQ2 and comparison to X-ray structure. (A) Schematic view of the DQ2 model. The backbones are represented as cyan (alpha chain) and blue (beta chain) ribbons. Secondary structure elements are represented as yellow arrows (beta strands) and red cylinders (alpha helices). The binding site for the peptide is given by the two N-terminal domains. (B) Comparison of secondary structure of the modelled (MODEL, this paper) and experimental (1S9V, PDB structure) structures of DQ2. Secondary structure was assigned with the DSSP program [33]. Secondary structures types are indicated according to the DSSP code (H: alpha helix; G: 3–10 helix; E: extended strand; T: H-bonded turn; S: bend; B: residue in isolated beta-bridge; none: coil). Sequence segments in bold highlight regions missing in the experimental structure (also indicated with dashes in the related secondary structure row). (C) Backbone superposition of modelled (blue) and experimental (red) structure of DQ2. The 105–112 region of B chain is absent in the X-ray structure. (D) Superposition of the modelled (blue) and experimental (red) amino acids involved in the interaction with the peptide in the main anchor positions. Alpha-I peptide is reported with by-atom colours.

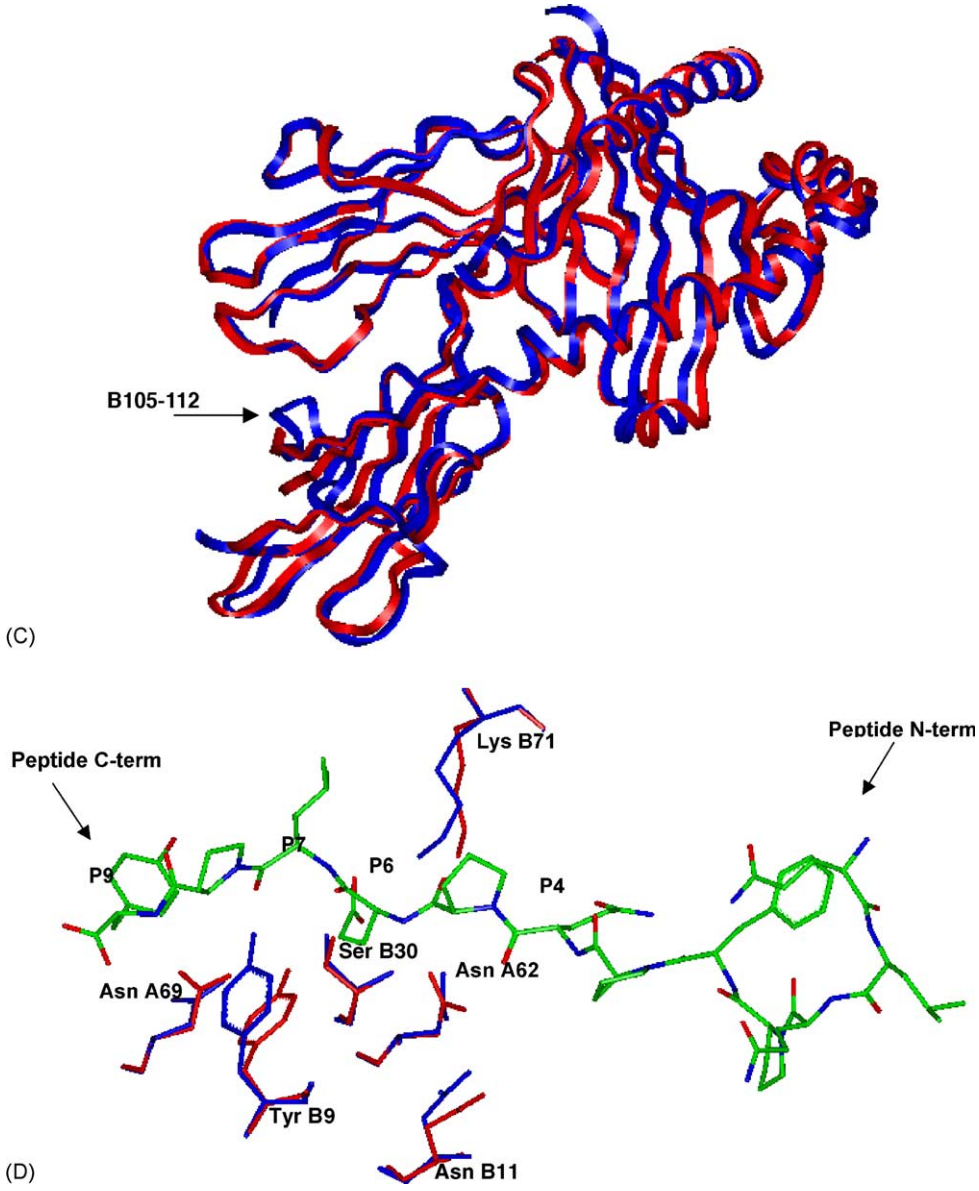


Fig. 2. (*Continued*).

with the neutral glycine do not affect in an evident way the total energy of interaction.

Concerning the alpha-I peptide, the analysis of the contributions to the overall energy of interaction seems to indicate that the Van der Waals interaction is unaffected by the substitution of glutamine with glutamate in p6, as expected by the structural similarity of the two side chains. On the contrary, the electrostatic energy is affected by the substitution and indicates a more favourable interaction for the glutamate side chain, thus suggesting that this side chain may interact in the binding site with positively charged groups of the DQ2. Therefore, we simulated the presence of a positively charged side chain in p6 of alpha-I by substituting the glutamine with a lysine; the calculated energy of interaction resulted more unfavourable than for the natural peptide, in agreement with the lower affinity experimentally observed. A Gln -> Gly substitution in p6 does not affect the interaction of the alpha-I peptide with DQ2, being the loss of Van der Waals interactions very similar to the gain of electrostatic energy.

In the simulations with alpha-II and its modified forms, the substitution of glutamine with glutamate in p4 improves the electrostatic energy of interaction and leaves quite similar the Van der Waals contribution, similarly to the alpha-I p6 substitutions. The Gln -> Lys substitution in p4 gives a little improvement of the electrostatic energy, while a glycine in p4 of alpha-II results in a loss of Van der Waals interaction. This should reflect differences between the p4 and p6 pockets in the DQ2-binding cavity.

We have also simulated the effects of substitution in p5 position of alpha-II (62–72). This position is considered not relevant for the interaction with the HLA molecule, being the

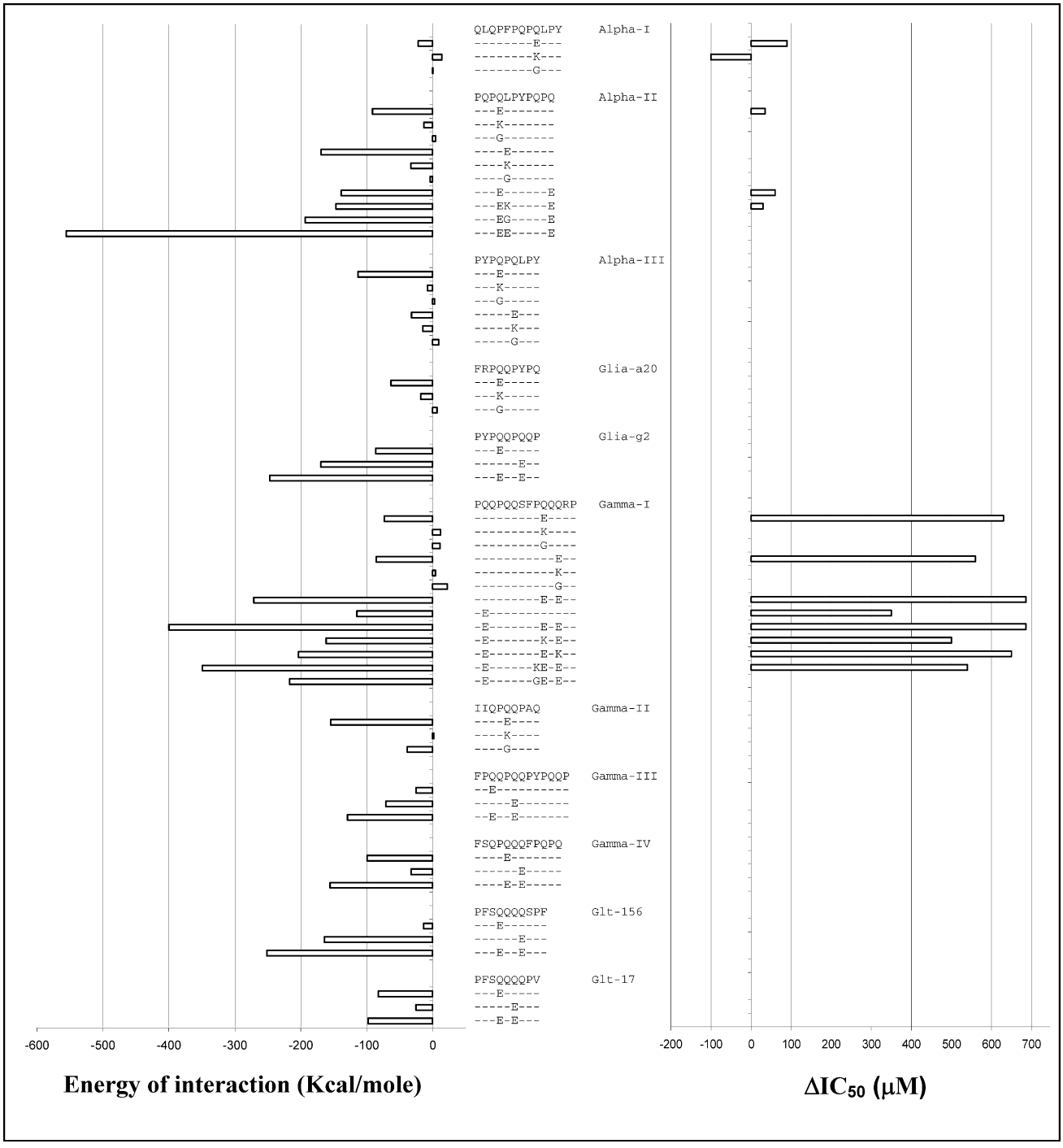


Fig. 3. Interaction peptide–DQ2. Left side: bars represent the difference of energy of interaction (in kcal/mol) in comparison to the natural peptide, computed as described under Section 2. The energy of interaction is given by the sum of the Van der Waals energy and electrostatic energy contribution. For each peptide, amino acids modifications have been simulated in different anchor positions: alpha-I, p6; alpha-II, p4, p5, and p11; alpha-III, p4 and p6; glia-a20, p4; glia-g2, p4 and p7; gamma-I, p-2, p6, p7, and p9; gamma-II, p5; gamma-III, p1 and p4; gamma-IV, p4 and p6; glt-156, p4 and p7; glt-17, p4 and p6. Right side: for a visual comparison to the computed energies of interaction, we report experimental affinity data, adapted from published figures and tables, for alpha-I [6], alpha-II [6], gamma-I [4] peptides and some their modified forms. No experimental data are available for other modified forms and other peptides. Values refer to the difference of IC50 (wt peptide − modified peptide).

side chain of this amino acid oriented toward the T-cell receptor [1]. A very little difference is observed between the energies of interaction for the alpha-II peptides having (p4)E(p11)E and (p4)E(p5)K(p11)E substitutions. To investigate in more detail the effect of charge in p5 position, we have simulated the substitution in p5 without deamidation in other positions. A glutamate side chain in p5 gives a relevant contribution to the interaction with the DQ2, while the lysine

also improves the interaction at a minor extent, and a glycine leaves unchanged the energy.

We simulated deamidation and other substitutions in alpha-III peptide at p4 and p6 positions. The deamidation of both glutamines improves the energy of interaction, while the substitution with glycine determines subtle variations of the energy. The substitution with the positively charged lysine gives a little improvement of interaction when it is in p4, while a more evident loss of interaction occurs when lysine is in p6. These observations are in good agreement with the results of our simulations with alpha-I and alpha-II.

The interaction between DQ2 and gamma-I peptide with substitutions of glutamines in p-2, p7 and p9 was simulated by considering that glutamine in p-2 of gamma-I is subjected at deamidation by tTG [4]. In our simulations, the negatively charged glutamate improves the total energy interaction in any of the three positions by a strong increase of electrostatic energy, and such increase becomes much more than additive when it is simulated the double or the triple substitution. Substitutions with lysine in different positions and in combination with glutamine deamidations gave intermolecular energy values in agreement with the experimental observation of lower DQ2 affinity for peptides with lysines.

We simulated deamidation in glia-g20 peptide at p4 and p7 positions and in glt-17 at p4 and p6 positions. The deamidation of these glutamines in both peptides improves the energy of interaction, and, in particular, the electrostatic contributions. Indeed, the effect of the deamidation and other substitutions was simulated in gamma-II peptide at p5 position and in glia-a20 at p4 position. In our simulations, the negatively charged glutamate improves the total energy interaction. The Gln -> Lys substitution in p5 for gamma-II and in p4 for glia-a20 gives a improvement of the electrostatic energy, while a glycine in these positions does not affect the interaction of these peptides with DQ2.

These results are in agreement with the experimental data recently published [39].

We also simulated deamidation effects on the interaction of gamma-III, gamma-IV and glt-156 with DQ2. For these peptides, it was experimentally shown their activity as DQ2-restricted T-cell epitopes and the tTG deamidation [37,38], but no study was performed to verify the effects of deamidation and the peptide position into the binding site. By similarity to the position of the other peptides, we found the optimal position for these peptides. Deamidation of the glutamines known to be subjected to this reaction by means of tTG determines an improvement of the electrostatic contribution to the energy of interaction, and double substitution improves the interaction much more.

3.3. Evaluation of H-bond interactions between DQ2 and peptides

To investigate in more detail the interaction of DQ2 with peptides, we analyzed the putative H-bonds in the peptide–DQ2 complexes, looking for differences related to the substitutions. The results of such analysis are reported in Table 1. The substitution of side chains in the anchor positions affects only the H-bonds involving the substituted side chain, without effects on the other side chains. Any substitution of glutamine with glutamate improves the H-bond network, by increasing the number of putative H-bonds, in many cases made by the negatively charged glutamate as acceptor and a positively charged side chain as donor, or at least by improving the geometry of the putative H-bond, which means lower distances and angles more close to 180°. The effect of substitutions with glycine is not reported in the table, because in any case its simple side chain, i.e. a hydrogen atom, reduces the potentiality to make H-bonds. Concerning the substitutions with lysine, in most cases they seem to decrease the number of the putative H-bonds and/or the quality in terms of geometric parameters, but in the case of alpha-I p6K and alpha-III p6K substitutions, a higher number of putative H-bonds are recognized, although their quality seems to be poor. Therefore, for these peptides, a complete evaluation of the contribution of substituted lysines to the formation of H-bonds is a very difficult endeavour.

4. Discussion

4.1. Comparison to published experimental structure of DQ2

We modelled the DQ2 structure and deposited it in the Protein Data Bank (PDB code: 1NBN) before the publication and availability of an experimental structure of DQ2 [15]. The comparison between the predicted model and the experimental structure confirms that the homology modelling procedure has been correctly applied. In fact, as shown in Fig. 2B and C, the superposition of chain backbones and secondary structures are very good. The few differences in the secondary structure assigned are mainly due to ‘‘G’’ (3–10 helix) recognized as ‘‘H’’ (alpha helix) or ‘‘T’’ (turn). A more deep comparison of side chains in the key pocket confirms a good overlap at the backbone level, but some differences are evident for the side chains. The modelling of side chain conformations is considered a step of the homology modelling strategies which needs to be improved. In our case, the good superposition of the backbone atoms suggests that side chain superposition may be improved by single bond rotations and side chain mobility, depending on the availability of free space in the pocket. We also noted the presence of water molecules in the experimental structure, but not in the predicted model, in particular in the epitope-binding site, which may affect the side chain geometry.

4.2. Comparison to published experimental results about peptide–DQ2 interaction

We created models of peptide–DQ2 complexes by using as template the experimental complex of the insulin peptide

Table 1
Putative H-bonds in the peptide–DQ2 complexes

Peptide	H-bond donor	H-bond acceptor	Distance H-A	Angle D-H-A
Alpha-I P6Q (wt)	Asn A71:ND2	Gln (P6):OE1	2.0	159.8
Alpha-I P6E	Asn A64:ND2	Glu (P6):OE1	1.8	166.3
	Asn A14:ND2	Glu (P6):OE2	1.9	158.3
Alpha-I P6K	Lys (P6):NZ	Asn A14:OD1	2.0	146.5
	Lys (P6):NZ	Asn A64:O	1.8	156.0
	Lys (P6):NZ	Val A67:O	2.2	151.0
Alpha-II P4Q (wt)	Asn A64:ND2	Gln (P4):O	2.1	152.4
	Gln (P4):NE2	Gly B13:O	1.8	159.1
Alpha-II P4E	Glu (P4):N	Tyr A11:O	2.0	164.1
	Lys B71:NZ	Glu (P4):OE2	1.7	165.4
Alpha-II P4K	Lys (P4):NZ	Ala A9:O	2.1	133.4
	Lys (P4):NZ	Gly B13:O	1.7	165.5
Alpha-II P5L (wt)	None			
Alpha-II P5E	Lys B71:NZ	Glu (P5):OE2	1.7	163.4
Alpha-II P5K	Lys (P5):NZ	Thr A63:O	1.9	148.0
Alpha-III P6Q (wt)	Gln (P6):N	Asn A64:OD1	2.1	158.7
	Asn A71:ND2	Gln (P6):OE1	2.0	160.9
Alpha-III P6E	Asn A64:ND2	Glu (P6):OE1	2.0	151.4
	Asn A14:ND2	Glu (P6):OE2	2.0	149.0
Alpha-III P6K	Lys (P6):NZ	Asn A14:OD1	2.2	133.4
	Lys (P6):NZ	Asn A64:O	2.0	143.2
Alpha-III P4Q (wt)	Gln (P4):NE2	Gly B13:O	1.9	152.9
Alpha-III P4E	Ser B28:OG	Glu (P4):OE2	1.6	156.5
	Lys B71:NZ	Glu (P4):OE2	2.1	131.8
Alpha-III P4K	Lys (P4):N	Tyr A11:O	2.0	158.6
	Lys (P4):NZ	Ala A9:O	2.1	123.5
	Lys (P4):NZ	Gly B13:O	1.7	171.0
	Lys (P4):NZ	Cys B15:SG	2.1	165.2
Gamma-I P7Q (wt)	Lys B71:NZ	Gln (P7):OE1	1.8	172.6
Gamma-I P7E	Arg B70:NE	Glu (P7):OE2	1.9	159.2
	Arg B70:NH2	Glu (P7):OE2	2.0	150.4
	Lys B71:NZ	Glu (P7):OE1	1.6	160.7
Gamma-I P7K	None			
Gamma-I P-2Q (wt)	None			
Gamma-I P-2E	Arg B88:NE	Glu (P-2):OE2	1.8	157.8
	Arg B88:NH2	Glu (P-2):OE2	1.8	161.0
Gamma-I P9Q (wt)	None			
Gamma-I P9E	Tyr B9:OH	Glu (P9):OE1	1.7	174.4
	Arg A78:NH1	Glu (P9):OE2	1.9	140.3
	Glu (P9):N	Asn A71:OD1	2.1	171.5
Gamma-I P9K	Lys (P9):N	Asn A71:OD1	1.9	175.2
Gamma-II P5Q (wt)	None			
Gamma-II P5E	Arg A70:NE	Glu (P5):OE1	1.8	162.7
Gamma-II P5K	Lys (P5):NZ	Phe A60:O	2.1	147.2
Gamma-III P1Q (wt)	Gln (P1):NE2	Glu B86:OE1	1.8	173.0
Gamma-III P1E	Phe A56:N	Glu (P1):OE1	1.9	164.5
Gamma-III P4Q (wt)	Asn A64:ND2	Gln (P4):O	2.2	150.2
	Gln (P4):NE2	Tyr A11:O	2.2	165.9
	Gln (P4):NE2	Gly B13:O	1.9	168.6
Gamma-III P4E	Glu (P4):N	Tyr A11:O	2.1	175.6
	Lys B71:NZ	Glu (P4):OE2	1.6	169.4
Gamma-IV P4Q (wt)	Gln (P4):N	Tyr A11:O	1.9	162.4
Gamma-IV P4E	Glu (P4):N	Tyr A11:O	1.9	159.4
	Ser B28:OG	Glu (P4):OE2	1.5	158.9
	Lys B71:NZ	Glu (P4):OE1	1.7	173.8
Gamma-IV P6Q (wt)	Asn A71:ND2	Gln (P6):OE1	1.9	156.9
Gamma-IV P6E	Asn A71:ND2	Glu (P6):OE1	1.9	155.7
	Asn A14:ND2	Glu (P6):OE2	2.1	143.6
	Asn A64:ND2	Glu (P6):OE2	2.0	151.9

Table 1 (*Continued*)

Peptide	H-bond donor	H-bond acceptor	Distance H-A	Angle D-H-A
Glia-a20 P4Q (wt)	Asn A64:ND2	Gln (P4):O	2.0	157.3
	Gln (P4):NE2	Gly B13:O	1.9	169.1
Glia-a20 P4E	Lys B71:NZ	Glu (P4):OE2	1.7	165.5
	Asn A64:ND2	Glu (P4):O	2.2	166.7
Glia-a20 P4K	Lys (P4):NZ	Ser A10:O	2.0	144.6
	Lys (P4):NZ	Gly B13:O	1.8	156.8
Glia-g2 P4Q (wt)	Gln (P4):N	Tyr A11:O	1.9	173.5
	Asn A64:ND2	Gln (P4):O	2.1	160.6
	Gln (P4):NE2	Gly B13:O	1.9	155.4
Glia-g2 P4E	Glu (P4):N	Tyr A11:O	2.0	161.9
	Ser B28:OG	Glu (P4):OE2	1.6	160.2
	Lys B71:NZ	Glu (P4):OE1	1.9	148.3
Glia-g2 P7Q (wt)	Lys B71:NZ	Gln (P7):OE1	2.0	167.5
Glia-g2 P7E	Arg B70:NE	Glu (P7):OE1	1.8	156.9
	Arg B70:NH2	Glu (P7):OE1	1.8	155.1
	Lys B71:NZ	Glu (P7):OE2	1.6	174.3
Glt-17 P4Q (wt)	Gln (P4):N	Tyr A11:O	1.9	166.8
Glt-17 P4E	Glu (P4):N	Tyr A11:O	2.0	159.0
	Lys B71:NZ	Glu (P4):OE1	1.7	163.7
Glt-17 P6Q (wt)	Asn A71:ND2	Gln (P6):OE1	2.0	164.4
Glt-17 P6E	Asn A71:ND2	Glu (P6):OE1	1.9	159.1
	Asn A14:ND2	Glu (P6):OE2	2.0	155.7
	Asn A64:ND2	Glu (P6):OE2	2.0	149.9
Glt-156 P4Q (wt)	Gln (P4):N	Tyr A11:O	1.9	172.9
	Gln (P4):NE2	Gly B13:O	1.8	158.7
Glt-156 P4E	Glu (P4):N	Tyr A11:O	2.0	171.7
	Lys B71:NZ	Glu (P4):OE1	1.6	164.1
	Ser B28:OG	Glu (P4):OE2	1.7	153.9
Glt-156 P7Q (wt)	Lys B71:NZ	Gln (P7):OE1	1.8	165.1
Glt-156 P7E	Arg B70:NE	Glu (P7):OE1	1.8	162.4
	Arg B70:NH2	Glu (P7):OE1	2.0	149.9
	Lys B71:NZ	Glu (P7):OE2	1.7	158.3

The H-bond donor and acceptor columns refer to the atom and amino acid in the peptide (identified by the anchor position in parentheses) or in DQ2 (A or B chain in the amino acid number). In order to consider only the stronger H-bond interactions, we reported only putative H-bonds with 2.2 Å of cut-off distance (see Section 2).

with DQ8. The substitution of the insulin side chains in order to simulate the appropriate gluten peptide sequence was used as strategy to model the peptide in the DQ2 ligand-binding site. Although it can be observed that more rigorous computational docking methods can be applied to create ligand–protein complexes, we preferred this strategy because it allowed us to locate specific peptide residues in the right anchor positions, according to the known mechanism of interaction of gluten peptides with DQ2 [1,2]. It should be also noted that the energies of interaction were computed for in vacuo models, without calculations of desolvation effects, and their values are heavily dependent on the dielectric constant used. Therefore, these calculated values cannot be considered the real values of interaction energies. For these reasons, in our study we used these values for a comparison among a number of complex models, but not as absolute measures of energy of interaction. By evaluating the energy of interaction and the putative H-bond formation, we observed that simulated substitutions improved or not the interaction peptide–DQ2, in excellent agreement with published experimental results. In fact, Arentz-Hansen et al. [6,37] identified different gluten peptides acting as DQ2-restricted T-cell epitopes and observed an increase of affinity after deamidation of Gln65 in the alpha 9 (57–68) peptide, and Gln65 in the alpha 2 (62–75) peptide, corresponding to p6 and p4 positions of alpha-I and alpha-II peptides, respectively, in a complex analysis based on truncated peptide variants. Moreover, these Authors have experimentally observed that the insertion of a lysine in p5 of deamidated alpha-II peptide leaves roughly unchanged the DQ2 binding. Quarsten et al. [4] verified by experimental methods the increase of affinity of gamma-I peptides deamidated at p-2, p7 and p9 anchor positions, as well as they demonstrated the importance of Lys-β71 for the peptide binding by using mutated DQ2. Vader et al. [39] have observed an increase of affinity after deamidation of glia-g2 to p4 and p7 positions, of glia-a20 at p4 position, of gamma-II at p5 position and of glt-17 to p4 and p6 positions. By a different approach, our results confirm that in different peptides the negatively charged side chain of glutamate is preferred in p4, p6, and p7 anchor positions, in comparison to the glutamine side chain, and at the same time the presence of positive charges in p6 and p7 of the peptide reduces the interaction with DQ2. The

negatively charged side chain of glutamate seems also to improve the interaction with DQ2 when it is in p-2, p1, p5 and p9, while it is assumed a preference of p9 position for hydrophobic and aromatic side chains, and no preference for the p5 position [1]. In addition, our study was extended on peptides for which no experimental data exists concerning the effects of deamidation on the interaction with DQ2. Our results indicate that the glutamine–glutamate substitution can improve the interaction DQ2–peptide for these peptides also, and we have suggested the assignment of anchor positions to the peptide residues. In general, our simulations are in very good agreement to the available experimental findings. In addition, more peptides and deamidation or substitutions have been simulated, and this work opens the opportunity to verify, at least at theoretical level, the effects of other peptide modifications.

4.3. Sequence analysis of DQ2 chains and differences with other alleles

We analyzed in more detail the molecular bases of the effects of the substitutions in the known anchor positions. The presence of many side chains with positive charge and few with negative charge characterizes the peptide-binding site in DQ2 (see Fig. 4). This observation may explain a generic preference of DQ2 to interact with peptides having negatively charged side chains in any position, also in p5 and p9 where it is not reported a preference for negatively charged amino acids. In particular, the region encompassing p4 and p7 positions is very close to the positively charged Lys-β71 of DQ2. We observed also that Lys-β71 is surrounded by two other positive charges, Arg-β70, which is also in the peptide-binding site, and Arg-β72, out of the binding site and not involved in the interaction with the peptide, but in contact with Glu-β46 and Glu-β69, thus it is probably involved in electrostatic interactions stabilizing the chain conformation. Arg-β70 and Lys-β71 are exposed on the surface of the peptide-binding site near the p4 and p7 site, so they can interact with the negative charge of glutamate of the peptide and improve the electrostatic energy of interaction; on the contrary, a lysine side chain in the peptide at each of these positions can create a repulsive electrostatic effect. This explanation is also supported by the evaluation of the energy of interaction for each peptide–DQ2 complex (see Fig. 3).

In addition to the comparison with DQ8 side chains reported in Fig. 1, we compared the amino acid sequence of DQ2 chains to all the DQ sequences available in public databases with particular attention to the charged amino acids involved into the peptide-binding site (the multiple alignment is not shown). The positively charged region Arg70-Lys71-Arg72 present in the beta chain of DQ2 (DQB1 *0201 allele) was found conserved only in the DQ2B1 *020x alleles, whereas in all others known alleles one or more positive charges are substituted by the neutral Thr, Gly or Ala, or by the negatively charged Glu or Asp. In more details, in the DQB1 *030x alleles it is present the Arg-Thr-Arg segment, mutated in Gly-Thr-Arg for the DQB1 *0308 and Glu-Asp-Arg in DQB1 *0306 allele. The Glu-Asp-Arg sequence is also in DQB1 *040x and DQB1 *0504, while the other DQB1 *050x present the segment Gly-Ala-Arg. In DQB1 *06xx the segment is Arg-Thr-Arg or Gly-Thr-Arg. In conclusion, Arg-β72 is conserved in all the alleles, while Lys-β71 is present only in the DQB1 *020x alleles, in some others it is substituted by the negatively charged side chain of aspartate. Moreover, the positively charged Arg-β70 is conserved in DQB1 *020x and some DQB1 *030x and DQB1 *060x alleles, but in all other alleles it is substituted, in many cases by the negatively charged side chain of glutamate. These negative charges may determine a

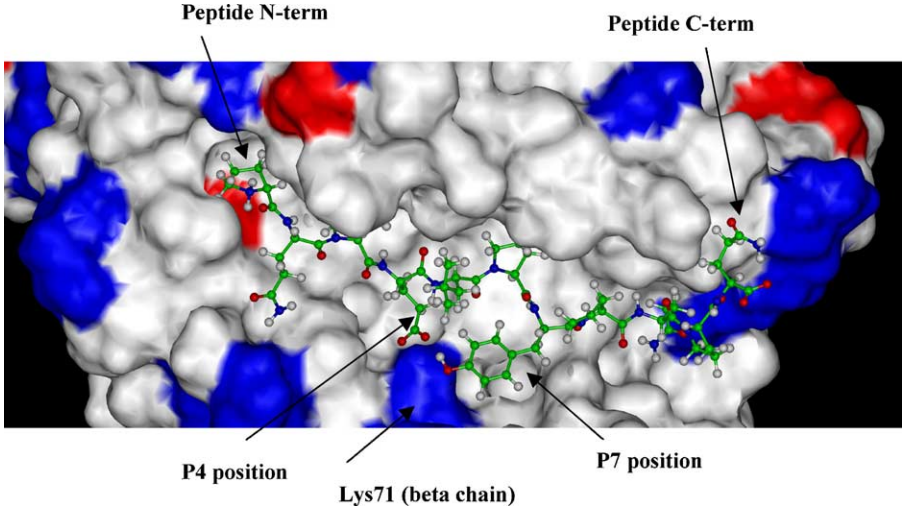


Fig. 4. Detailed view of the surface of the binding site of DQ2. The alpha II is shown with ball-and-stick representation and standard atom colours. The DQ2 surface is coloured to evidence the presence of negatively charged (red) and positively charged side chains (blue). The N-terminal (NH_3^+) of the peptide is near a negative charge, while the C-terminal (COO^-) of the peptide is near a positively charged region. The positions of p4 and p7 side chains in the peptide are evidenced, as also the Lys71 of the DQ2 beta chain.

repulsive electrostatic effect with the negative charges of peptides. These observations suggest that the DQ2 alleles found in celiac patients may specifically interact with gliadin peptides, in particular if deamidated, by means of electrostatic interactions between the positive region of Arg-β70-Lys-β71 and its surrounding and the negative charged peptides. Other HLA-DQ may have a lower affinity for such deamidated gliadin peptides, due to the lack of these positive charges in the binding site. These observations are also in agreement with the work of Quarsten et al. [4] which substituted the Lys-β71 with a threonine (similarly to the substitution appearing between DQB1*030x and DQB1*06xx alleles) and observed that original and mutated (Gln -> Glu) peptides were active as stimulating antigens for the T-cell recognition when presented by DQ2, but the stimulation resulted 100-fold poorer with the mutated DQ2.

To investigate the role of other charged residues, we observed that Arg-β77 is present only in DQB1*020x and DQB1*050x alleles, while a threonine is present in all other alleles. This is also an interesting observation, because Arg-β77 is very close to the p-2 position, considered not important for DQ2-peptide recognition. Another positive side chain, Arg-β88, very close to the p-2 position, is conserved in all alleles. This positively charged residue might explain the improvement of interaction when we substituted a glutamine with a glutamate in p-2 of gamma-I peptide (see previous paragraph).

We observed also that Glu-β69, which is very close the Arg-β72, is conserved in all alleles, thus confirming the possibility that the interaction Glu-β69-Arg-β72 may be important to stabilize the chain conformation. The other glutamate very close to Arg-β72, i.e. Glu-β46, is present only in DQB1*020x alleles, while in all other alleles this position is occupied by a valine.

Few charged amino acid of alpha chain are located in the peptide-binding site. Arg-α78 results conserved in all alleles, on the contrary Arg-α55 shows interesting mutations. In this region, DQA1*3xx have two arginines, DQA1*01xx have two glycines, but all other alleles have only one arginine and a gap for the second amino acid. Therefore DQA1*3xx alleles, i.e. the alpha chain of DQ8, have two positive charges in a region of the peptide-binding site where it is generally located the N-terminal side of the peptide, while DQ2 has only one arginine.

5. Conclusions

For many years, studies at the molecular level of the interaction between DQ2 and gluten peptides were based on molecular simulations and modelling of DQ2. In order to improve the previous molecular modelling results by using the most recent and reliable approach, we created a model of DQ2 based on a better homologous model and by creating a full atom model instead of simple side chain substitution, and deposited the model in PDB to make it available to the scientific community for further investigation. Then, we simulated the interaction with different peptides, with many amino acid substitutions, in order to verify the agreement of simulation results with experimental reports about the binding of gluten peptides and their modifications. The agreement between experimental and simulation results appears very good and, as a first conclusion, this work demonstrated how simulations can be successful applied to analyse and understand, at molecular level, experimental results of peptide binding, in particular for molecular event related to celiac disease.

We also give an interpretation for the role of Lys-β71 as well as of other charged residues in the binding site and peculiar of DQ2. Our conclusions about the involvement of Lys-β71 in the interaction with the peptide p4 and p7 anchor sites are in good agreement with the experimental analyses concerning the peptide-binding regions. In addition, our simulations give details about the effects of amino acid substitutions, as deamidation of glutamines and substitution of negatively charged glutamate with positively charged lysine residues, which were experimentally tested for the binding to DQ2, but still not investigated for structural characterization by experimental methods.

In conclusion, our work demonstrates how molecular modelling and binding simulations can be applied to understand at molecular level the initial recognition events related to the celiac disease, in particular the binding of gluten peptides to DQ2 and how it may be affected by peptide modifications. This work opens the perspective of further investigations about other peptides related to celiac disease as well as to understand the effects of other peptide modifications, with the aim to obtain more structural and functional information about the peptide recognition by DQ2 as well as other HLA molecules.

Acknowledgements

Ph.D. fellowship of Dr. Susan Costantini is supported by E.U.

References

[1] L.M. Sollid, Molecular basis of celiac disease, Ann. Rev. Immunol. 18 (2000) 53–81.
[2] L.M. Sollid, Celiac disease: dissecting a complex inflammatory disorder, Nat. Rev. Immunol. 2 (2002) 647–655.
[3] D. Schuppan, Current concepts of celiac disease pathogenesis, Gastroenterology 119 (2000) 234–242.
[4] H. Quarsten, O. Molberg, L. Fugger, S.N. McAdam, L.M. Sollid, HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope, Eur. J. Immunol. 29 (1999) 2506–2514.
[5] A.K. Moustakas, Y. van de Wal, J. Routsias, Y.M. Kooy, P. van Veelen, J.W. Drijfhout, F. Koning, G.K. Papadopoulos, Structure of celiac disease-associated HLA-DQ8 and non-associated HLA-DQ9 alleles in complex with two disease-specific epitopes, Int. Immunol. 12 (2000) 1157–1166.

[6] H. Arentz-Hansen, R. Korner, O. Molberg, H. Quarsten, W. Vader, Y.M. Kooy, K.E. Lundin, F. Koning, P. Roepstorff, L.M. Sollid, S.N. McAdam, The intestinal T cell response to a-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase, J. Exp. Med. 191 (2000) 603–612.
[7] O. Molberg, S.N. McAdam, R. Korner, H. Quarsten, C. Kristiansen, L. Madsen, L. Fugger, H. Scott, O. Noren, P. Roepstorff, K.E.A. Lundin, H. Sjostrom, L.M. Sollid, Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease, Nat. Med. 4 (1998) 713–717.
[8] Y. van de Wal, Y. Kooy, P. van Veelen, S. Pena, L. Mearin, G. Papadopoulos, F. Koning, Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity, J. Immunol. 161 (1998) 1585–1588.
[9] H. Sjostrom, K.E. Lundin, O. Molberg, R. Korner, S.N. McAdam, D. Anthonsen, H. Quarsten, O. Noren, P. Roepstorff, E. Thorsby, L.M. Sollid, Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition, Scand. J. Immunol. 48 (1998) 111–115.
[10] Y. van de Wal, Y.M.C. Kooy, J.W. Drijfhout, R. Amons, G.K. Papadopoulos, F. Koning, Unique peptide binding characteristics of the disease-associated DQ(a1*0501, b1*0201) vs the non disease associated DQ(a1*0201,b1*0202) molecule, Immunogenetics 46 (1997) 484–492.
[11] R. Kumar, A.L. Eastwood, M.L. Brown, G.W. Laurie, Human genome search in celiac disease: mutated gliadin T-cell-like epitope in two human proteins promotes T-cell activation, J. Mol. Biol. 319 (2002) 593–602.
[12] F. Vartdal, B.H. Johansen, T. Friede, C.J. Thorpe, S. Stefanovic, J.E. Eriksen, K. Sletten, E. Thorsby, H.-G. Rammensee, L.M. Sollid, The peptide binding motif of the disease associated HLA-DQ (a1*0501, b1*0201) molecule, Eur. J. Immunol. 26 (1996) 2764–2772.
[13] G.K. Papadopoulos, C. Wijmenga, F. Koning, Interplay between genetics and the environment in the development of celiac disease: perspectives for a healthy life, J. Clin. Invest. 108 (2001) 1261–1266.
[14] P.E. Bourne, K.J. Addess, W.F. Bluhm, L. Chen, N. Deshpande, Z. Feng, W. Fleri, R. Green, J.C. Merino-Ott, W. Townsend-Merino, H. Weissig, J. Westbrook, H.M. Berman, The distribution and query systems of the RCSB Protein Data Bank, Nucl. Acids Res. 32 (2004) D223–D225.
[15] C.Y. Kim, H. Quarsten, E. Bergseng, C. Khosla, L.M. Sollid, Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease, PNAS 101 (12) (2004) 4175–4179.
[16] A.C.R. Martin, M.W. MacArthur, J.M. Thornton, Assessment of comparative modelling in CASP2, Proteins (Suppl. 1) (1997) 14–28.
[17] R. Rodriguez, G. Chinea, N. Lopez, T. Pons, G. Vriend, Homology modelling, model and software evaluation: three related resources, Bioinformatics 14 (1988) 523–528.
[18] D.R. Westhead, J.M. Thornton, Protein structure prediction, Curr. Opin. Biotech. 9 (1998) 383–389.
[19] A. Tramontano, Homology modelling with low sequence identity, Methods 14 (1998) 293–300.
[20] C. Caporale, C. Caruso, A. Facchiano, M. Nobile, L. Leonardi, L. Bertini, G. Colonna, V. Buonocore, Probing the modelled structure of wheatwin1 by controlled proteolysis and sequence analysis of unfractionated digestion mixtures, Proteins 36 (1999) 192–204.
[21] A.M. Facchiano, P. Stiuso, M.L. Chiusano, M. Caraglia, G. Giuberti, M. Marra, A. Abruzzese, G. Colonna, Homology modelling of the human eukaryotic initiation factor 5A (eIF-5A), Protein Eng. 14 (2001) 881–890.
[22] C. Caporale, A. Facchiano, L. Bestini, L. Leopardi, G. Chiosi, V. Buonocore, C. Caruso, Comparing the modelled structures of PR-4 proteins from wheat, J. Mol. Model (Online) 9 (2003) 9–15.
[23] A. Marabotti, S. D'Auria, M. Rossi, A.M. Facchiano, Theoretical model of the three-dimensional structure of a sugar binding protein from Pyrococcus horikoshii: structural analysis and sugar binding simulations, Biochem. J. 280 (2004) 677–684.
[24] G. Scapigliati, S. Costantini, G. Colonna, A. Facchiano, F. Buonocore, P. Bossù, J.W. Holland, C.J. Secombes, Modelling of fish interleukin 1 and its receptor, Dev. Comp. Immunol. 28 (2004) 429–441.
[25] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[26] K.H. Lee, K.W. Wucherpfennig, D.C. Wiley, Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes, Nat. Immunol. 2 (2001) 501–507.
[27] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[28] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.
[29] C.A. Orengo, A.D. Michie, S. Jones, D.T. Jones, M.B. Swindells, J.M. Thornton, CATH—a hierarchic classification of protein domain structures, Structure 5 (1997) 1093–1108.
[30] F.M.G. Pearl, D. Lee, J.E. Bray, I. Sillitoe, A.E. Todd, A.P. Harrison, J.M. Thornton, C.A. Orengo, Assigning genomic sequences to CATH, Nucl. Acids Res. 28 (2000) 277–282.
[31] A.G. Murzin, S.E. Brenner, T. Hubbard, C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, J. Mol. Biol. 247 (1995) 536–540.
[32] L. Lo Conte, S.E. Brenner, T.J.P. Hubbard, C. Chothia, A. Murzin, SCOP database in 2002: refinements accommodate structural genomics, Nucl. Acid Res. 30 (2002) 264–267.
[33] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[34] S.J. Hubbard, S.F. Campbell, J.M. Thornton, Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors, J. Mol. Biol. 220 (1991) 507–530.
[35] I.K. McDonald, J.M. Thornton, Satisfying hydrogen bonding potential in proteins, J. Mol. Biol. 238 (1994) 777–793.
[36] S. Sarkhel, G.R. Desiraju, N–H···O, O–H···O, and C–H···O hydrogen bonds in protein–ligand complexes: strong and weak interactions in molecular recognition, Proteins 54 (2004) 247–259.
[37] H. Arentz-Hansen, S.N. McAdam, O. Molberg, B. Fleckenstein, K.E. Lundin, T.J. Jorgensen, G. Jung, P. Roepstorff, L.M. Sollid, Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues, Gastroenterology 123 (2002) 803–809.
[38] W. Vader, Y. Kooy, P. Van Veelen, A. De Ru, D. Harris, W. Benckhuijsen, S. Pena, L. Mearin, J.W. Drijfhout, F. Koning, The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides, Gastroenterology 122 (2002) 1729–1737.
[39] W. Vader, D. Stepniak, Y. Kooy, L. Mearin, A. Thompson, J.J. van Rood, L. Spaenij, F. Koning, The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses, PNAS 100 (21) (2003) 12390–12395.