Paliakasis-K; **Routsias-JG**; Petratos-K; Ouzounis-C; Kokkinidis-M; Papadopoulos-GK *Νέα δομικά χαρακτηριστικά των ανθρωπίνων αντιγόνων ιστοσυμβατότητας* ***HLA-DQ προκύπτουν απο την μοντελοποίηση τους με βάση την δημοσιευμένη*** *δομή των μορίων HLA-DR.* J-Struct-Biol. 1996 Sep-Oct; 117(2): 145-63.

Προηγούμενες μελέτες έχουν υποδείξει την γενετική συσχέτιση του τόπου DQ στην επιρρέπεια ή στην αντίσταση σε ορισμένες αυτοάνοσες παθήσεις όπως ο διβήτης τύπου 1, η κατά πλάκας σκλήρυνση και η κοιλιοκάκη. Στη μελέτη αυτή παρουσιάζονται τα δομικά χαρακτηριστικά των μορίων HLA-DQ, όπως αυτά προέκυψαν από μοριακή μοντελοποίηση (λογισμικό Insight και Discover, Biosym Technologies) με βάση την ήδη γνωστή δομή του ομολόγου προς αυτά μορίου HLA-DR. Τα αποκλειστικά χαρακτηριστικά των HLA-DQ μορίων περιλαμβάνουν: (α) μια πολυμορφική $1^{η}$ θέση αγκίστρωσης στην κοιλότητα δέσμευσης αντιγόνου, (β) μια πολυμορφική περιοχή διμερισμού (β49-56) και (γ) σε πολλά αλλήλια ένα βρόχο Arg-Gly-Asp (β167-169), ο οποίος ενδεχομένως να ενέχεται σε μία διαδικασία κυτταρικής προσκόλλησης. Η περιοχή διμερισμού α2β2 και το τμήμα δέσμευσης του CD4 εμφανίζονται σχεδόν ταυτόσημα με τις αντίστοιχες περιοχές στη δομή του HLA-DR1. Τα HLA-DQ διαφοροποιούνται από τα HLA-DR μόρια στις θέσεις αγκίστρωσης 1,2 και/ή 5, γεγονός που καταδεικνύει τις διαφορές τους ως προς την επιλογή πεπτιδικών αντιγονικών θραυσμάτων. Επιπλέον ο πολυμορφισμός στην περιοχή ομοδιμερισμού (β49-56) προσδιορίζει τον βαθμό ευκολίας για διμερισμό κάτω από την επίδραση του Τ υποδοχέα των λεμφοκυττάρων. Με δεδομένη την αναγκαιότητα ομοδιμερισμού για την ενεργοποίηση των Τ-λεμφοκυττάρων και των αντιγονο-παρουσιαστικών κυττάρων, η δυνατότητα διμερισμού των DQ αλληλίων αντικατοπτρίζει την ευκολία ή δυσκολία ενεργοποίησης των δύο παραπάνω κυτταρικών τύπων. Επιπρόσθετα, ο βρόχος RGD δίνει τη δυνατότητα κυτταρικής προσκόλλησης στα αλλήλια DQ που τον κατέχουν.

Τα δομικά χαρακτηριστικά που παρουσιάζονται στη μελέτη αυτή, δίνουν πολύτιμες πληροφορίες για τους πιθανούς μηχανισμούς με τους οποίους καθορίζονται οι μοναδικές ανοσολογικές ιδιότητες των HLA-DQ μορίων.

JOURNAL OF STRUCTURAL BIOLOGY **117,** 145–163 (1996)
ARTICLE NO. 0079

Novel Structural Features of the Human Histocompatibility Molecules HLA-DQ as Revealed by Modeling Based on the Published Structure of the Related Molecule HLA-DR[1]

KOSTAS PALIAKASIS[2]

Institute of Molecular Biology and Biotechnology, Research Center of Crete, Heraklion, Crete, Greece

JOHN ROUTSIAS[2]

Laboratory of Immunology, Department of Internal Medicine, University of Ioannina Medical School, GR-451 10 Ioannina, Greece

KYRIACOS PETRATOS

Institute of Molecular Biology and Biotechnology, Research Center of Crete, Heraklion, Crete, Greece

CHRISTOS OUZOUNIS

European Molecular Biology Laboratory, Meyerhoffstrasse 1, D-69012 Heidelberg, Germany

MICHAEL KOKKINIDIS

Institute of Molecular Biology and Biotechnology, Research Center of Crete, Heraklion, Crete, Greece

AND

GEORGE K. PAPADOPOULOS[3]

Laboratory of Immunology, Department of Internal Medicine, University of Ioannina Medical School, GR-451 10 Ioannina, Greece

Received January 15, 1996, and in revised form June 11, 1996

Structural modeling of the HLA-DQ molecules, a group of human histocompatibility antigens linked to autoimmune diseases and immunosuppression-based on the structure of the homologous molecule DR1, has revealed an overall shape typical of the class II histocompatibility molecules, yet with several novel features. These are unique to HLA-DQ and include: (1) an antigen-binding groove with a polymorphic first pocket and anchoring in the second and/or fifth pocket, (2) a polymorphic β49–56 dimerization patch, and (3) in many alleles a prominent Arg-Gly-Asp loop (β167–169), probably involved in cell adhesion, as it exhibits an architecture similar to identical sequences involved in such function. The α2β2 dimerisation domain and the CD4-binding region are nearly identical to their counterparts in the structure of HLA-DR1. The significance of the few substitutions in the CD-4 binding region remains to be evaluated. The polymorphic first antigen-binding pocket and the anchoring in the second and/or fifth pocket point to differences in antigenic fragment selection compared to HLA-DR antigens, while the polymorphism in the β49–56 homodimerization patch implies either ease of spontaneous or T lymphocyte receptor-induced homodimerization or difficulty in the latter. As homodimerization appears to be an obligatatory intermediate in the activation of cognate DQ-restricted T lymphocytes and DQ-bearing antigen-presenting cells, the dimerization properties of DQ allels signify the respective ease or difficulty of activation of these two cell types. The RGD loop confers cell adhesion possibilities to those DQ allels that possess it, yet its putative

[1] This work was presented in part at the winter meeting of the Hellenic Biochemical and Biophysical Society, Athens, Greece, January 1994 and the 13th Immunology and Diabetes Workshop, Montvillagerne, France, May 1994.

[2] The contribution of the first two authors is equal.

[3] To whom correspondence should be addressed. Fax: 30 651 45944.

1047-8477/96 $18.00

ligand cannot be defined at present. These features are suggestive of the probable mechanisms through which some of the unique immunological properties of the HLA-DQ molecules are effected.

INTRODUCTION

The class II region of the human major histocompatibility complex (MHC)[4] consists of several genes and pseudogenes grouped together in a number of distinct gene loci (Trowsdale and Campbell, 1992; Campbell and Trowsdale, 1993). Within this region the loci of DP, DQ, and DR have received the most attention. Each locus contains one gene for an α chain and one or more genes for a β chain. The genes for DR are expressed constitutively in antigen-presenting cells (macrophages, B lymphocytes, and dendritic cells) while the other two gene loci can be induced by cytokines and certain viruses. The mature class II MHC molecules are integral membrane proteins on the surface of antigen presenting cells, with the α and β chains non-covalently linked to each other and mostly containing a tightly bound ($K_D \sim \mu M$–nM) antigenic peptide (Jardetzky *et al.*, 1990; Sette *et al.*, 1989). Complementation between chains of different HLA class II loci has also been observed (Lotteau *et al.*, 1987). The two chains are synthesized together with an invariant (Ii) chain that is an obligatory chaperone of the αβ heterodimer until the complex reaches the endosome compartment (Germain, 1994). There, the acidic environment leads to the dissociation of the invariant chain from the heterodimer and the trapping of antigenic peptides from internalized and degraded protein antigens. The αβ MHC class II heterodimer containing by the antigenic peptide is then transported to the cell membrane where it can be recognized by cognate $CD4^+$ T helper lymphocytes. These lymphocyte emerge from the thymus organ, with an antigen receptor that recognizes peptide antigen presented (i.e., bound) by a specific class II histocompatibility molecule of the organism, a property known as MHC restriction (Male *et al.*, 1991). This recognition coupled with another signal or recognition event leads to activation of helper T lymphocytes and the antigen-presenting cells bearing the DQ molecules (Janeway, 1992).

Recently the three-dimensional structure of the extracellular domains of the human MHC Class II molecule DR1 (DRA/DRB1*0101) became known to 3.3 Å resolution (Brown *et al.*, 1993), and a higher-resolution structure (2.75 Å) of the same DR1 molecule complexed to the influenza virus hemagglutinin peptide HA306–318 was subsequently published (Stern *et al.*, 1994). The chief characteristics of the DR1 molecule in the antigenic peptide-binding region (α1β1 domain) are: a molecular shape very similar to that of MHC class I molecules, an antigen-binding groove that is open at both ends allowing long (9–22 residues) peptides to bind, and five residue-binding pockets of which the first is the most prominent. One of the surprising findings in the elucidation of the structure of DR1 has been the documentation of the DR homodimer of αβ heterodimers in the crystallographic unit cell (pH of crystallization 4–6) (Brown *et al.*, 1993); This homodimer of heterodimers is also known in the complex of DR1 with the staphylococcal superantigens SEB (Jardetzky *et al.*, 1994), and TSST-1 (Kim *et al.*, 1994), and more recently in the complexes of DR3 with the CLIP peptide (Ghosh *et al.*, 1995b), and DR1, and a peptide from HLA-A8 (Stern, 1995). This homodimerization has been hypothesized to occur *in vivo* via cross-linking by identical T lymphocyte receptors on the cognate T lymphocyte, recognizing two identical DR molecules bearing the same antigenic peptide (Brown *et al.*, 1993).

A number of population studies have linked the DQ locus to susceptibility or resistance to certain autoimmune diseases such as type 1 diabetes, multiple sclerosis, coeliac disease, etc. (see Altman *et al.*, 1991, for a review). The linkage to DQ is more highly correlated to these diseases than linkage to DR or DP (respectively centromeric and telomeric to DQ). An epistatic function of the DQ locus over the DR locus was found in another study, without any apparent explanation (Hirayama *et al.*, 1987). A third study provided evidence that DQ molecules participate in immunosuppression, by mechanisms that are not clearly understood (Salgame *et al.*, 1991), while other work has claimed that soluble DQ molecules are generated by alternative splicing (Briata *et al.*, 1989), a finding that has yet to be verified by another laboratory.

No crystal structures of HLA-DQ molecules have been reported thus far. In order to gain insight into the putative structural differences between DQ and DR molecules, which are probably responsible for the functional differences reported, we have taken advantage of the extensive similarity (65–70% identity in amino acid positions) between them and modeled the three-dimensional structure of the former using the published structure of DR1. In addition to the extensive identity in amino acid positions, the DQ and DR molecules have many functional similarities that suggest similar structures for nearly identical functions: they both bind

[4] Abbreviations used: HLA, human leukocyte antigens; MHC, major histocompatibility complex; RGD, arginine-glycine-aspartate; rms, root mean square.

antigenic peptides that are more than 12 amino acids in length (Gedde-Dahl *et al.*, 1993; Jones *et al.*, 1994), with similar kinetics and pH-dependency of binding (Johansen *et al.*, 1994), and with the antigenic peptide having at least one chief anchor residue (Falk *et al.*, 1994; Verreck *et al.*, 1994). Furthermore, DQ- and DR-restricted T lymphocyte clones both use the coreceptor molecule CD4 (Fossum *et al.*, 1993). There are a number of examples in the literature (HIV-1 protease (Weber *et al.*, 1989), human islet β cell glucokinase (St. Charles *et l.*, 1994)) in which this functional criterion has been used successfully to model proteins with similar function, yet only 30–35% identity in amino acid sequences (Wlodawer *et al.*, 1989). In fact, in a recent retrospective study of the modeling of the three-dimensional structure of several proteins, it was noted that such modeling was very successful in cases of more than 40% amino acid identity between proteins of similar function (Srinivasan and Blundell, 1993). In a study just published we showed that the modeled structural features of select HLA-DQ alleles segregate according to the susceptibility or resistance conferred by them to type 1 diabetes mellitus, an autoimmune disease (Routsias and Papadopoulos, 1995). Here we analyze the structural characteristics of DQ molecules and provide an account of the structural basis for the differences in immune function between HLA-DQ molecules and their DR and DP counterparts.

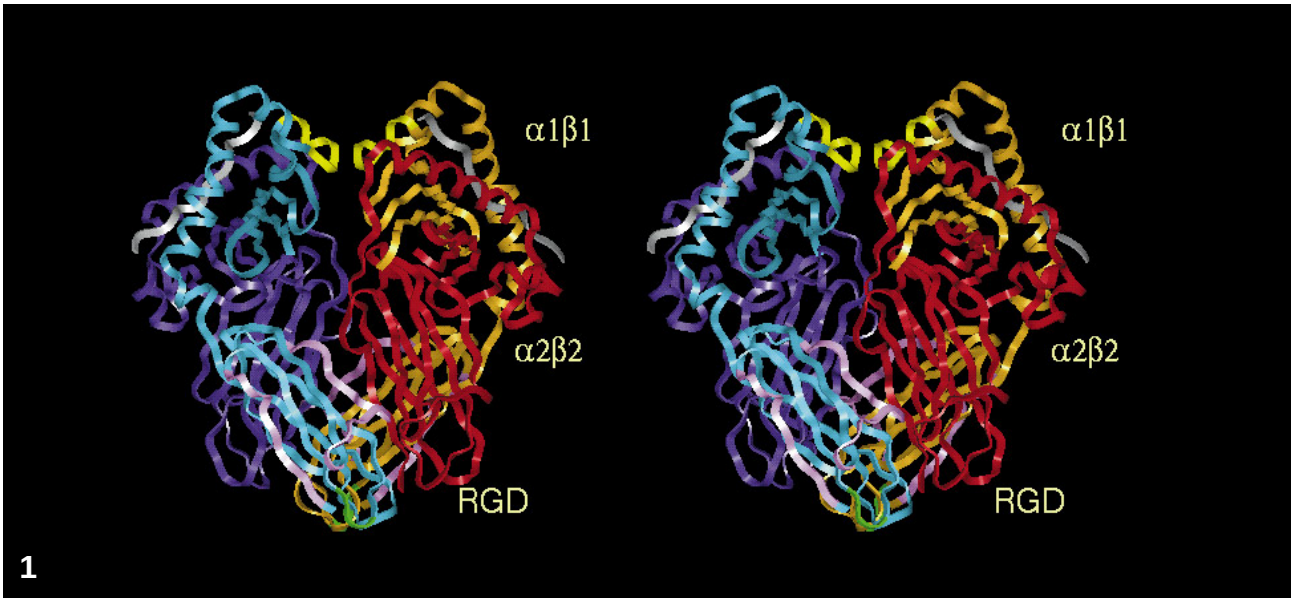
MATERIALS AND METHODS

The coordinates of the HLA-DR1 heterodimer (Brown *et al.*, 1993) and the homodimer of heterodimers complexed to the HA306–318 peptide (Stern *et al.*, 1994) were kindly provided by Professor Don C. Wiley. They include the positions of residues 3–182 of each of the two α chains and 3–190 of each of the two β chains, as well as the coordinates of the two trapped HA306–318 peptides. Initially the structure of DQA1*0101/DQB1*0501 (entries HA25-human and HB22-human in SwissProt (Bairock and Boeckman, 1991)) was built, followed by models of DQA1*0102/DQB1*0602, DQA1*0501/;DQB1*0301, DQA1*0501/DQB1*0302, DQA1*0301/DQB1*0301, and DQA1*0301/DQB1*0302. The first model building was performed on an Evans and Sutherland work station using Sybil (Tripos Assoc., St. Louis). HLA-DR and -DQ sequences were obtained from Kabat *et al.* (1991) and Marsh and Bodmer (1995) and supplemented by a search through the EMBL data bank. The alignment of DQ and DR molecules was taken from Brown *et al.* (1988). Subsequent modeling work was performed on a Silicon Graphics Iris workstation using the program Insight II, version 2.3.1 (Biosym Technologies, San Diego, CA). There was one insertion in position DQα9, which was matched to position DRα6. The previous amino acids (positions 6, 7, and 8) in DQα occupied positions 3, 4, and 5, respectively, in the DR1 structure, in order to accommodate the insertion and preserve all interactions of identical amino acids from position DQα9 onward. There were no amino acid substitutions in the sequence that would be expected to change dramtically the local conformation of any allelic molecule. The amino acid conformations were automatically chosen from a library of rotamers provided by the porogram, using the most suitable rotamer for each case. Any automatic energy minimization was in place after the replacement of each substituted amino acid residue. Interactions between identical amino acids in equivalent positioins were preserved. The ionization state of amino acid side chains was that at pH 5.0, the approximate pH of the endosomal compartment and also the average pH for the crystallization of DR1 (Brown *et al.*, 1993; Stern *et al.*, 1994; Jardetzky *et al.*, 1994; Ghosh *et al.*, 1995). Because MHC class II molecules have been shown to be very stable *in vitro* when complexed with antigenic peptides, we have incorporated a 13-residue polyalanine as the bound peptide of the DQ molecules modeled, using the coordinates of the bound hemagglutinin peptide HA306–318 in its complex with DR1. Energy minimization was accomplished by the conjugate gradient method. The minimization procedure went through 200 cycles. In the first 30 to 40 cycles of energy minimization the energy of the molecule dropped dramatically, while thereafter there were only very small changes in the energy values. The force field used included electrostatic terms in interactions up to 16 Å. No account was taken of the water molecules, as the high-resolution (2.75 Å) structure of HLA-DR1 on which modeling was based identified only one water molecule within the antigen-binding groove. Water molecules seem to play an important role as bridges between polar side chains of amino acids in the antigen-binding groove of the histocompatibility molecule and amide and carbonyl groups of the bound antigenic peptide, as revealed in the high-resolution structures of H-2K^b (Matsumura *et al.*, 1992; Fremont *et al.*, 1995), HLA-DR1 with an HLA-A8 peptide (Stern, 1995), and I-E^k with a hemoglobin peptide (Fremont *et al.*, 1996). For comparison purposes the published structure of DR1 was subject to the same minimization procedure yielding an average root mean square (rms) deviation for all C_α atoms of 0.52 Å and for all atoms of 0.76 Å. Pockets in the antigen-binding groove were tentatively identified as depressions in the van der Waals representation of the α1β1 domain. Graphical representations of the modeled molecules were obtained on the Silicon Graphics Iris work station using the program Insight II.

The hydrophobic profile of the various β49-56 and of eight or longer residue stretches of water-soluble proteins of known three-dimensional structure was calculated according to Engelman *et al.* (1986). The search for membrane proteins containing an RGD sequence in their extracellular domain was performed through the EMBL data bank. The coordinates of the modeled structure of HLA-DQA1*0101/DQB1*0501 will be deposited at the Brookhaven data bank. These coordinates will in the meantime be provided to interested researchers.

RESULTS

The overall modeled structure of the DQ molecules is very similar to that of DR1 (Fig. 1). Indeed, in most cases the very same amino acid residues in identical positions interact with one another in order to produce very similar biological effects, such as antigen binding, α2β2 dimerization, and CD4 binding (Fig. 2). The mean rms deviation between DR1 and DQA1*0101/B1*0501 is 0.51 Å for all C_α atoms, 0.58 Å for all backbone atoms, 0.77 Å for all heavy atoms, and 1.11 Å for all toms. Comparison of the coordinates of DR1 with those of other DQ molecules generated gave similar results. The many fine structural features of DQ molecules in the antigen-binding groove, the homodimerization patch of β49–56, the CD4-binding region, and the presence of a novel RGD loop, which are probably in part responsible for the distinct properties of these molecules, are analyzed below.



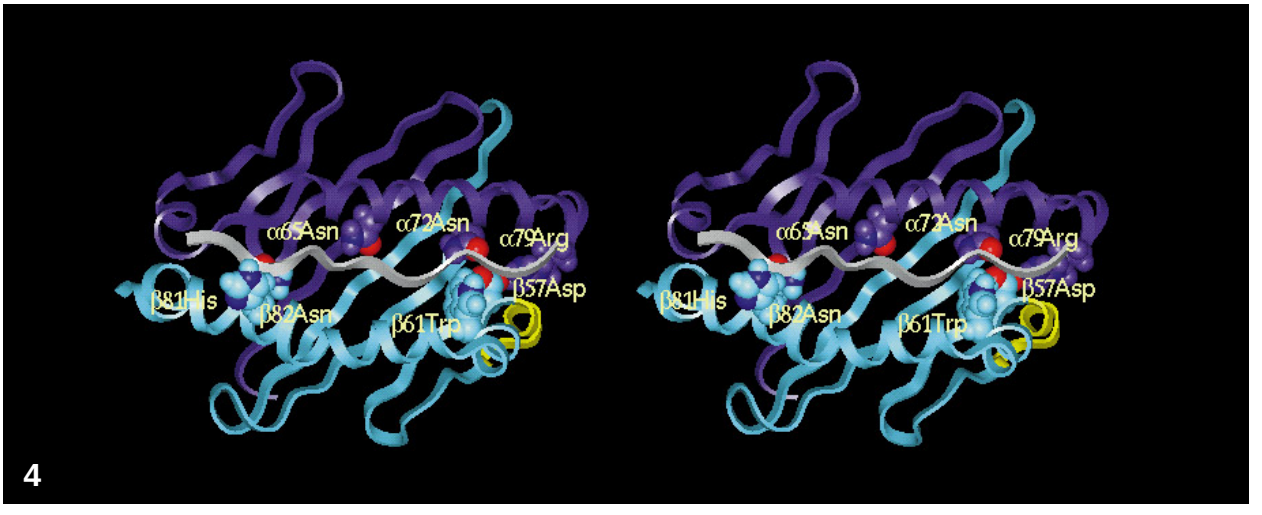
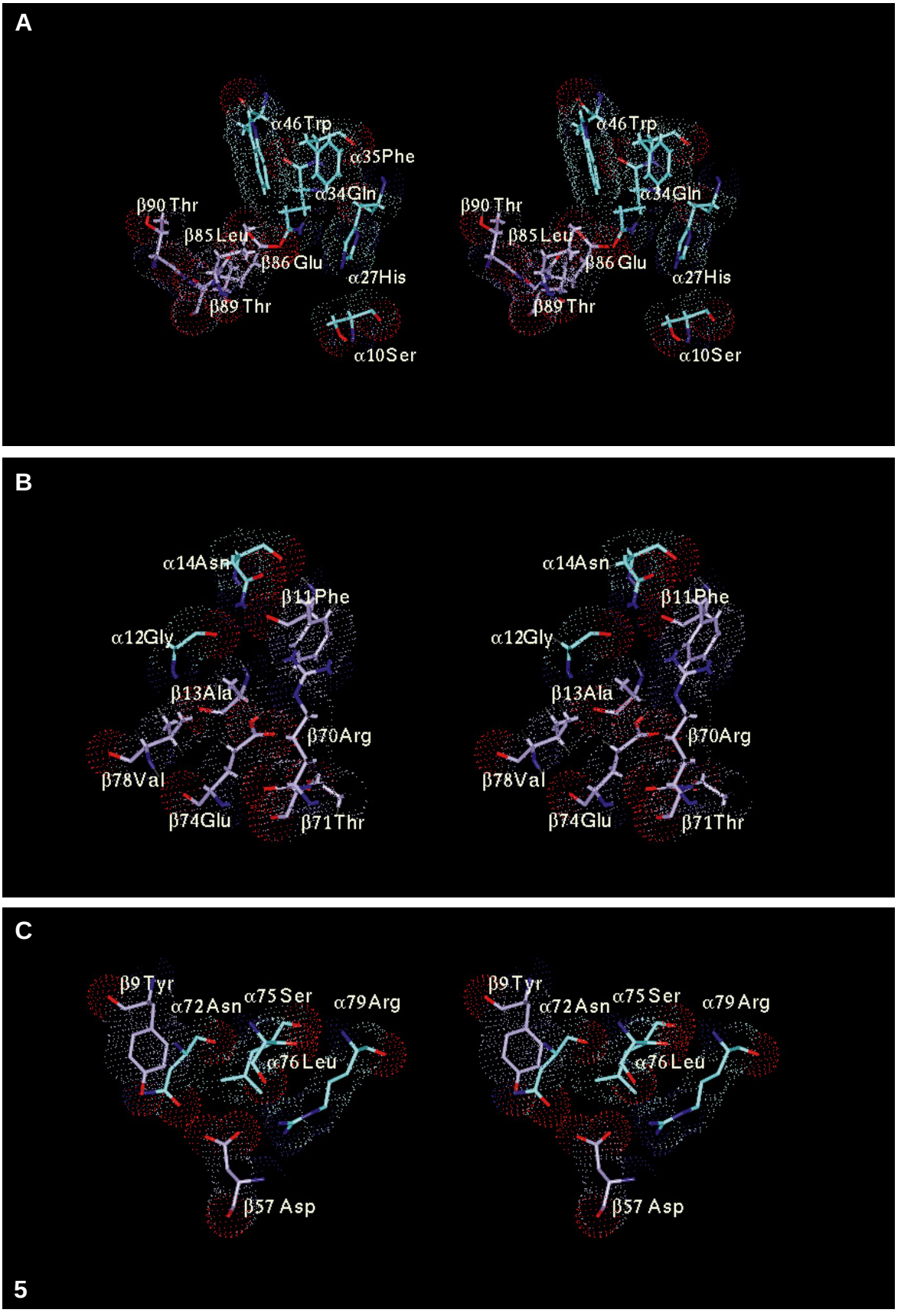


FIG. 1. Stereo view of the overall structure of the extracellular part of a typical DQ molecule (allele DQA1*0501/B1*0302) as a homodimer of two αβ heterodimers, obtained after amino acid substitution on the high-resolution coordinates of DR1 (Stern *et al.*, 1994) and energy minimization. The DQ molecule on the left has the β chain (light blue) toward the viewer and the α chain (purple) away from the viewer, while the DQ molecule on the right has the α chain (red) toward the viewer and the β chain (orange) away from the viewer. The two gray ribbons in the α1β1 domains are virtual antigenic peptide backbones from the actual coordinates of HA306–318 in its complex with DR1. The antigen-binding α1β1 domain of each heterodimer consists of the two α-helices and eight β sheets that bind the antigenic peptide. The two B49–56 dimerization patches are shown in yellow at the top. The α2β2 domain is below the α1β1 domain and contains the CD4-binding region, the RGD loop, and several crucial contact residues for homodimerization (see Figs. 6 and 8 for greater detail). The CD4-binding region on the β chains (β134–148 and β180–189) is in light pink. There are two identical CD4-binding regions per homodimer of heterodimers (one in front of the viewer and one on the opposite side of the homodimer). The two β167–169 RGD loops are in green at the bottom of the figure. The plasma membrane of the cell bearing the DQ molecule starts about 10 residues below the RGD loop and is parallel to the horizontal.

FIG. 4. Stereo view of the backbone tracing of the foreign antigen binding α1β1 domain of the modeled structure of the DQA1*0501/DQB1*0301 molecule, obtained after energy minimization (α chain in purple, β chain in light blue), as viewed from the top downwards. The backbone trace of the mock antigenic polyalanine peptide (modeled initially from the coordinates of the influenza hemagglutinin peptide HA306–318) (light gray) is shown for orientation pruposes. The amino terminus of the mock antigenic peptide is on the left. The amino terminus of the α chain is at the bottom, while that of the β chain is at the top of the figure. Residues α65Asn, α72Asn, and α79Arg and β57Asp, β61Trp, β81His, and β82Asn (all in space-filling representation, with their oxygen atoms in red and nitrogen atoms in deep blue) are invariant in all DR and DQ molecules and point toward the peptide-binding groove, forming hydrogen bonds to amide and carbonyl groups of the antigenic peptide. The first two residues of the mock antigenic peptide most probably form hydrogen bonds with the amide nitrogen and carbonyl oxygen of residues α54 and α56, as in DR1 (Stern *et al.*, 1994). For orientation purposes the homodimerization sequence β49–56 is shown in yellow.


A
α46Trp
α35Phe
α34Gln
β90 Thr
β85 Leu
β86 Glu
α27His
β89 Thr
α10Ser
α46Trp
α34Gln
β90 Thr
β85 Leu
β86 Glu
α27His
β89 Thr
α10Ser
B
α14Asn
β11Phe
α12Gly
β13Ala
β70Arg
β78Val
β74Glu
β71Thr
α14Asn
β11Phe
α12Gly
β13Ala
β70Arg
β78Val
β74Glu
β71Thr
C
β9 Tyr
α72 Asn
α75 Ser
α79 Arg
α76 Leu
β57 Asp
β9 Tyr
α72 Asn
α75 Ser
α79 Arg
α76 Leu
β57 Asp

5

The Antigen Binding Groove

The α1β1 domain of the DQ molecule exhibits the basic fold of histocompatibility molecules, i.e., a "floor" of eight β-pleated sheets bounded on two sides by two antiparallel α-helices (Fig. 1, and Bjorkman *et al.*, 1987; Fremont *et al.*, 1992; Brown *et al.*, 1993). There is a 55–60% identity in amino acid positions in the α1β1 domain of any DQ molecule, compared to DR1. Within the various DQ molecules (14 allelic α and 26 allelic β chains) there is a 75–80% identity in this same domain (Figs. 2 and 3, and Brown *et al.*, 1988). Despite the polymorphism in both α and β chains of DQ alleles, all of the invariant residues deemed crucial for forming hydrogen bonds to the foreign peptide's backbone in the case of DR1 are found in the same positions in DQ (Stern *et al.*, 1994, and Fig. 4).Specifically, the residues α65Asn, α72Asn, α79Arg (62, 69, and 76 for DRα) and β61Trp, β81His, and β82Asn are found in very similar orientations to those as in DR1 and forming hydrogen bonds with the same atoms from the mock antigenic peptide. In addition, the peptide groups of amino acids α54 and α56 of DQ are oriented in a very similar manner to those of DR (equivalent positions 51 and 53, respectively), enabling the former to form hydrogen bonds with peptide groups at the amino terminus of the trapped antigenic peptide, as in DR1 (Stern *et al.*, 1994). The hydrogen bond formed in DR1 with the antigenic peptide by residue α9Gln is not observed in DQ as the equivalent position α12 is occupied by glycine. Likewise, position β71 in DQ is polymorphic (mostly threonine, but also arginine or glutamate or aspartate) which might be able to form hydrogen bonds with the antigenic peptide, as β71Arg does in DR1, depending on the peptide's orientation. In DQ molecules with aspartate in position β57, this residue forms a salt bridge to the apposed α79Arg as in DR1. Furthermore, two hydrogen bonds are formed between these two residues and proximal amide and carbonyl groups of the polyalanine mock antigenic peptide. Indeed whenever the aspartate is present in β57 the mock antigenic peptide runs over the salt bridge formed by this aspartate and the apposed arginine. The bound peptide makes slight movements in the groove, relative to the initial position of HA306–318 in DR1, especially within the frame of the second pocket, which is very deep in DQ molecules (see below).

The antigen-binding groove of the modeled DQ molecules fulfills the "pocket" motif of this region for class II MHC molecules (Figs. 4 and 5, and Brown *et al.*, 1993; Stern *et al.*, 1994). All the depressions in the antigen-binding groove may not belong to pockets, so the number of pockets and the residues constituting each pocket may slightly vary depending on the nature of the antigenic peptide bound (Stern *et al.*, 1994). The first pocket in modeled DQ molecules is outlined by residues α10, α27, α34, α35, α46, β85, β86, β89, and β90, just as in DR1 (Figs. 2, 3, and 5A, and Stern *et al.*, 1994). In the different DQ molecules this first pocket may appear essentially as one of four different combinations. When these combinations are looked at as pairs, the amphiphillic and the hydrophilic pair, the only difference within each pair is an α34Glu → Gln substitution. The residues contributed by the DQα chain are either identical to (35Phe, 46Trp) or smaller and more hydrophilic (α10Ile → Ser, α27Phe → His, α34Ile → Glu or Gln) than those of the DRα chain. The variability in the first pocket in the DQ molecules is determined by the residues contributed by the β chain, which can be either β85Val, β86Ala (or Gly in 3/9 cases), β89Gly, and β90Ile, or β85Leu, β86Glu, β89Thr, and β90Thr. In the former case we have a slightly larger and amphiphilic pocket (9 alleles), and in the latter a slightly smaller and more hydrophilic pocket that could have one positively charged and one or two negatively charged residues (10 alleles). There are 7 alleles with unknown sequences in this region (Fig. 3, and Marsh and Bodmer, 1995).

The second pocket appears as the most prominent in DQ molecules (Fig. 5B). It is defined by residues α12, α14, β11, β13, β70, β71, β74, and β78, which are the same residues that constitute the second pocket in DR1, except α14 and β11. The latter two residues (Asn and Phe, respectively) are actually the lower

FIG. 5. Stereo view of the first, second, and fifth pockets in the antigen-binding groove of the modeled DQ molecule (DQA1*0501/DQB1*0301). Van der Waals surfaces of the participating residues are indicated by dots. The residues from the α chain are shown in light blue and those from the β chain are shown in light purple. The oxygen atoms of the residues are in red and the nitrogen atoms in deep blue. Orientation of the pockets is identical to that of Fig. 4. (A) The first pocket of this DQ allele belongs to the hydrophilic variant. Note that in many DQA alleles position 34 may be occupied by a glutamate instead of a glutamine, making the pocket more hydrophilic. The amphiphilic variant of the first pocket has hydrophobic residues from the β chain, which are not as bulky, making the first pocket of such a variant less shallow. (B) The modeled second pocket of the antigen-binding groove for the same molecule. Note that nearly all other DQB alleles have Gly in position β13, making this pocket even deeper; furthermore, residues α12 and β78 are invariant in all DQ alleles, while there is extensive variability in residues β70, β71, and β74. (C) The modeled fifth pocket of DQA1*0501/DQB1*0301. There are polymorphisms in several positions (α72 and 73, β9 and 57), but the pocket can generally accommodate residues of some bulk (e.g., Leu, Glu). In DQB alleles with aspartate in position β57, instead of serine, valine, or alanine, the binding of acidic residues in this pocket is probably hindered. By contrast, all DQ alleles not having Asp at β57 do allow the binding of an acidic residue at the fifth pocket. An alanine or a valine at this position points downward and away from α79Arg (data not shown).

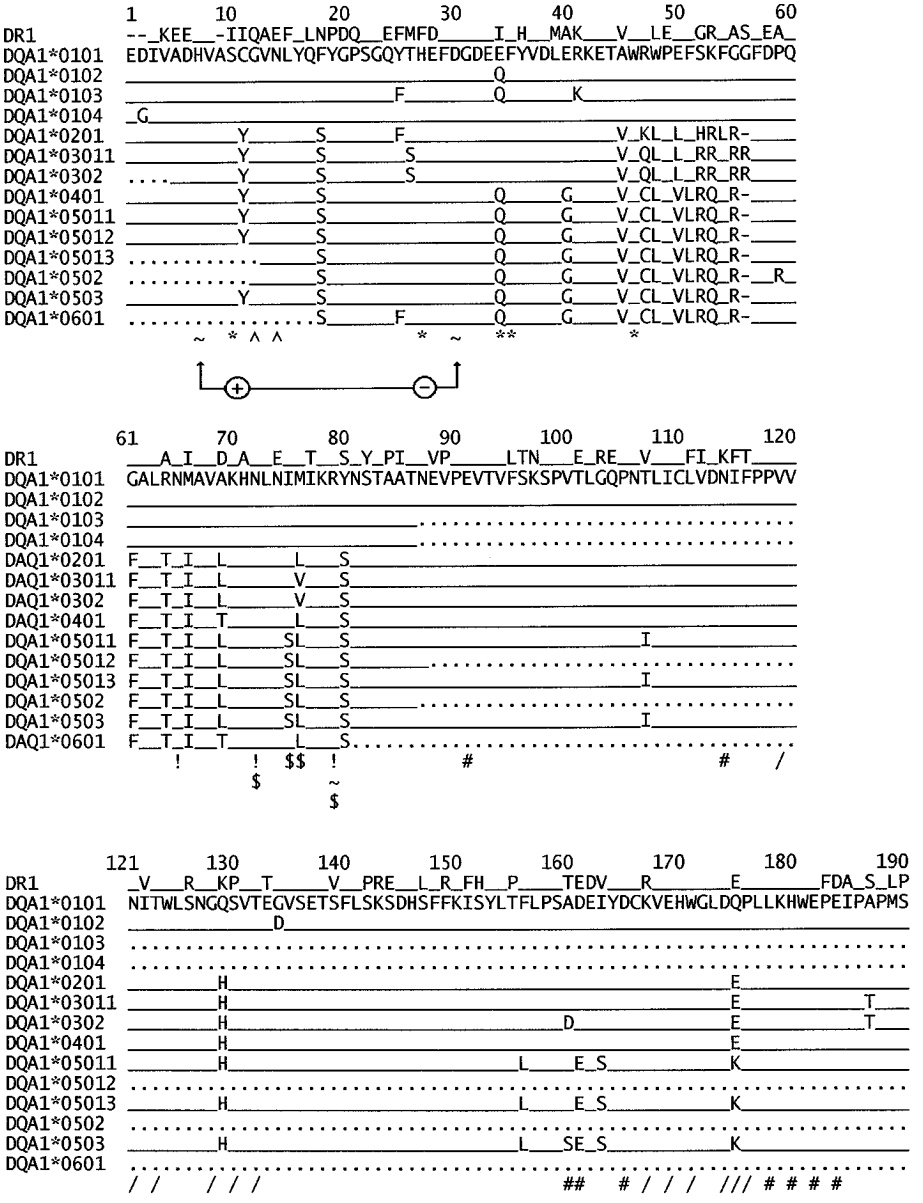


FIG. 2. Comparison of the amino acid sequences of DR1 and the known DQA alleles. There is ca. 55 and 65% identity in the α1 and α2 domains of DR1 and any DQA allele, respectively. Amino acids are in the one-letter code. DR1α has three amino acids less than DQα. The numbering in the α chain refers to DQ. Code: (_) identity; (-) no amino acid in corresponding position; (.) residue unknown; (~) intramolecular salt bridges (involved residues indicated in 1/2 cases by arrows below the ~ sign and a straight line linking the two residues and showing the charge of each in a circle; the second case involves α79Arg and wherever possible β57Asp); (!) intermolecular hydrogen bonds between HLA-DQ or -DR and the bound antigenic peptide; (*) residues forming the first pocket; (/\) residues forming the second pocket; ($) residues forming the fifth pocket; (/) residues forming the CD4-binding region; (#) interacting residues in the α2β2 domain stabilizing the homodimer.

boundaries of the pocket, since residues α12 and β13, originating from the β-sheet floor and pointing up toward the groove, are glycines (except in DQB1*0301, 0304, 06011, and 06012 that have an alanine at β13). The glycine and alanine residues in these positions are much smaller than the corresponding ones in DR1 (Gln and Phe, respectively). As such they enable this pocket to accommodate a bulky hydrophobic residue. The last four residues of the pocket are very polymorphic, yet no grouping of polymorphisms can be made into variants of the pocket with distinct physicochemical character. There are examples in the literature, however, where these top residues determine the nature of the amino acid bound in the second pocket because of electrostatic considerations (see Discussion below).

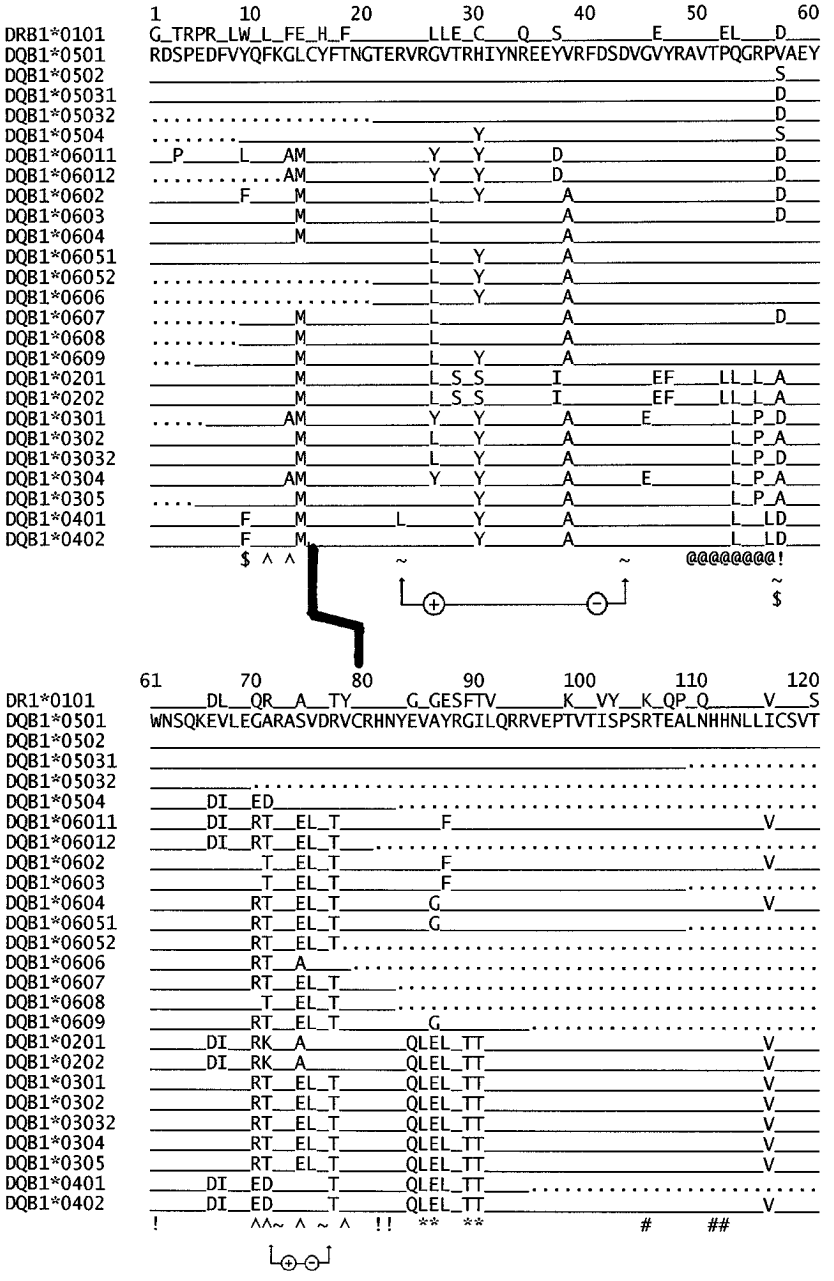


FIG. 3. Comparison of the amino acid sequences of DR1 and known DQB alleles. There is ca. 60 and 70% identity in the β1 and β2 domains of various DQB alleles compared to the respective domains of DR1. Code as in Fig. 2. The two cystenes (14 and 79) forming a disulfide bridge are shown as linked by a heavy line. (@) residues of the β49–56 dimerization patch; the β167–169 RGD tripeptide is in bold type.

The depressions in the remainder of the antigen-binding groove make for another three pockets at most, which correspond to the pockets made by DR1 in complex with HA306-318. Equivalently positioned residues participate in the formation of these depressions that we tentatively identify as pockets. Of these three depressions the last, corresponding to pocket 5, appears as prominent and would probably accommodate bulky hdrophobic residues, including the aromatic ones, due to the tyrosine (17/21 alleles), phenylalanine (3/21), or leucine (1/21) in position β9 instead of tryptophan of DR1 (Figs. 3 and 5C, and Stern *et al.*, 1994). Of the four residues contributed to this pocket by DQα, those in positions 75 and 76 are invariably hydrophobic (Ile or Ser for the former and Leu, Met, or Val for the latter). This pocket also contains the polymorphic residue β57 (Asp in many but also Val, Ser, or Ala in the other DQ alleles).

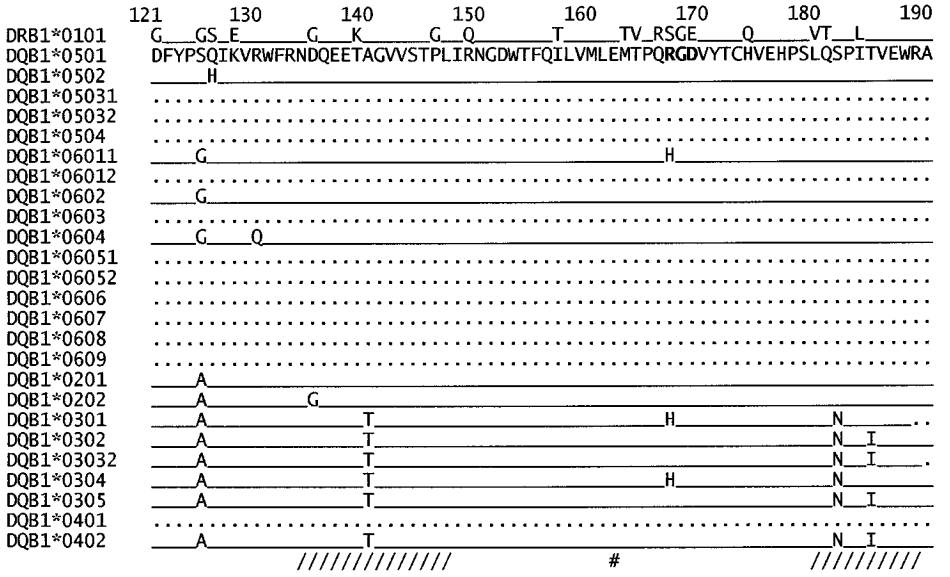


FIG. 3—*Continued*

Indeed, the character of the pocket changes dramatically, as revealed by the comparison of this pocket in DQ3.1 (A1* (0501/B1*0301) and DQ3.2 (A1*0501/B1*0302), obtained through modeling: in the latter, the β57Ala residue points downward and away from the apposed α79Arg, indicative of the stabilizing attraction between Arg and Asp that is now partly lost (data not shown). A similar shift is seen with DQA1*0101/B1*0501 where β57 is a valine (data not shown).

The three salt bridges that contribute to the stability of MHC class I and class II molecules in the α1α2 and α1β1 domains, respectively, are invariant in DQ molecules. These are formed between α7His (α5 in DR) and α30Asp (α27 in DR), β25Arg, and β43Asp, and β72Arg and β76Asp (Figs. 2 and 3, and Brown *et al.*, 1988). Because of an amino acid insertion in position 9 of the DQα chain, the distance between α7His and α30Asp in the DQ heterodimer is nearly double the corresponding distance in the DR1 and other DR molecules, thus diminishing the strength of this interaction.

α2β2 Dimerization Domain and the β49-56 Dimerization Patch

In the crystal structure of DR molecules and the modeled structure of their DQ homologues there is a major area involved in homodimerization within the α2β2 domain and a lesser one in the β1 domain (Figs. 1–4 and 6). We examined the amino acid substitutions in the dimerization patch of the α2β2 domain, i.e., $(\alpha 2)_A$ with $(\beta 2)_B$ and $(\beta 2)_A$ with $(\alpha 2)_B$, where A and B denote two different DQαβ heterodimers forming a homodimer. In the α chain there are two conservative substitutions (α161E → D, α184D → E), one semiconservative (α114K → N) and one nonconservative (α160T → A), while the remaining five amino acids involved in the stability of the homodimer of heterodimers are intact. Four of the five residues from the β chain involved in the α2β2 homodimer formation remain unchanged, while there is a conservative substitution, β105K → R (Figs. 3 and 6). The contacts between the various residues in the $(\alpha 2)_B$ and $(\beta 2)_A$ domains are very similar to those observed in the homodimer of DR1 heterodimers (Figs. 2, 3, and 6). There is also one interaction between four amino acid side chains of the two α2 domains involving two α91Glu and two α114Asn residues. All four side chains are located in the interior of the homodimer and oriented in a way to assure maximum hydrogen bonding and minimal repulsion between negatively charged glutamate side chains (Fig. 6). The importance of the residues in α114 has been questioned, however (Goodman *et al.*, 1995, and Discussion, below).

The second site of dimerization interaction in DR molecules $[(\beta 49–56)_A \ldots (\beta 49–56)_B]$ has a totally different sequence in the DQ counterpart (Fig. 3). The sequence is invariant in 138/141 DRB1, B3, and B5 alleles (Marsh and Bodmer, 1995). Of the four residues making the homodimer contact in DR, β49Ala, 51Thr, 52Glu, and 55Arg, the first two remain unchanged in all DQB alleles known to date. The glutamate residue of DRβ52 is changed to proline in 24 DQ alleles and to leucine in 2 alleles, while β455Arg is unchanged in 18 DQ alleles, but changed to leucine in 2 and to proline in 6 other alleles (Fig. 3). Consequently, the four salt bridges

TABLE I
Hydrophobicity Profile of the DQβ49-56 Dimerization Sequence

Protein	Sequence	Number of alleles	Free energy of transport to a hydrophilic environment (per residue) kcal/mole
	49 56		
DR	AVTELGRP	48	−11.5 (−1.34)
DQ			
1.	___PQ___	16	−10.4 (−1.30)
2.	___L__L_	2	+14.6 (+1.82)
3.	___P__P_	6	+8.6 (+1.08)
4.	___P___L	2	−0.5 (−0.06)

Note. Calculated from the hydrophobicity–hydrophilicity scale provided by Engelman *et al.* (1986).

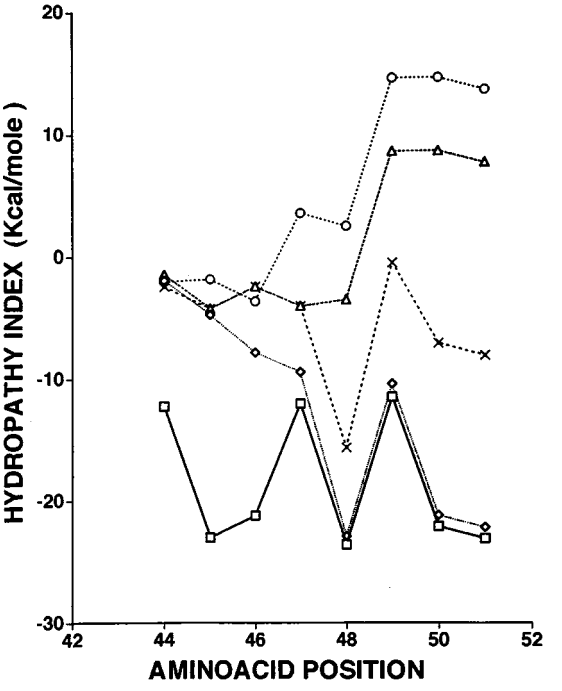


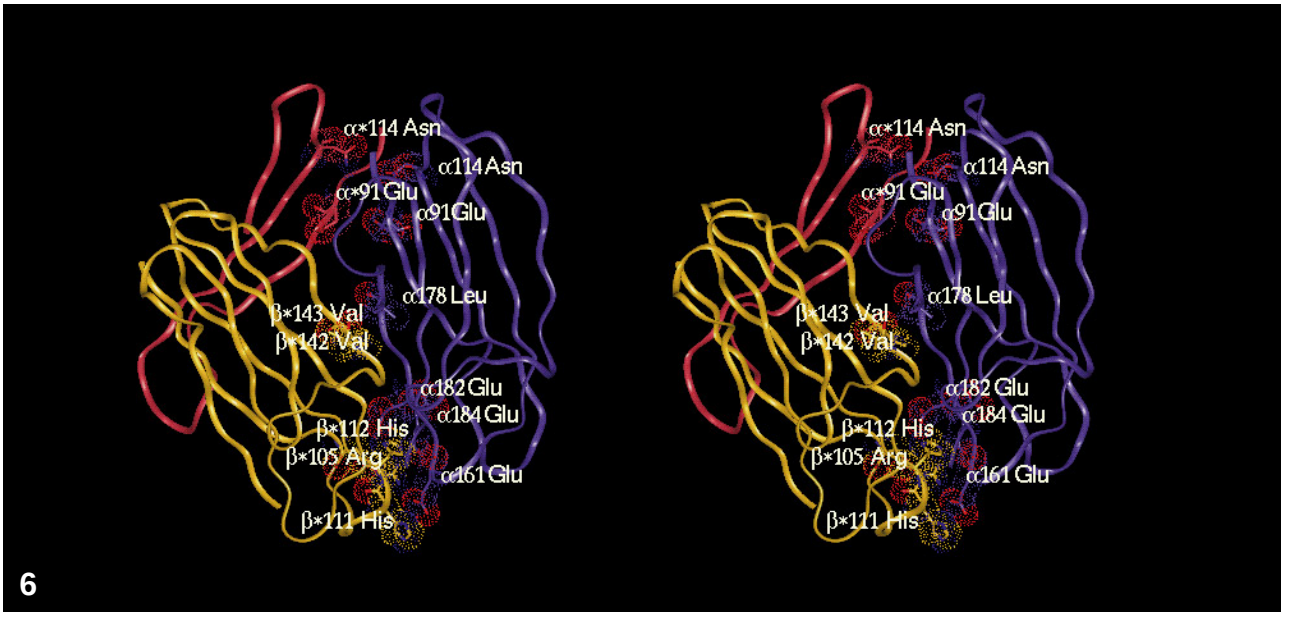
FIG. 7. Hydropathy plot of the amino acid sequences around the β49–56 dimerization domain. A window of eight amino acids were used. The hydropathy scale was from Engelman *et al.* (1986). The DQ allelic groupings are as shown in Table I. (□) DRβ (◇) DQβ group 1; (○) DQβ group 2; (△) DQβ group 3; (×) DQβ group 4. Some of the alleles in DQβ group 1 would have respectively values of −9.4, or −11.4 kcal/mole for position 50, as they have respectively valine or serine instead of aspartate in position 57; likewise, the respective values for position 51 would be −10.4 and −12.4 kcal/mole. In DQβ group 3, the aspartate instead of alanine in position β57 would make the hydrophobicity value for positions 50 and 51 be −1.6 and −2.6 kcal/mole, respectively.

between the two β52Glu and the two apposed β55Arg that promote the stability of the DR homodimers in this region are nonexistent in DQ molecules. Furthermore, in the 8 latter cases of proline or leucine at position β55, a leucine is present in β53, adding to the hydrophobicity of the patch (Table 1). In fact, a plot of the hydrophobicity values using an eight-amino-acid stretch window, and starting from DQβ44 up to and including DQβ58, shows that in all DQ alleles the sequence β49–56 is the most hydrophobic compared to neighboring octapeptide sequences (Fig. 7). Comparison of the hydrophobicity of this β49–56 patch with other peptide stretches known to participate in dimer or oligomer formation (e.g., hemoglobin S contact tetrapeptide, NF-κB p50 dimer contact amino acids in all but one cases, TNFα trimer contact regions, etc.) shows that the hydrophobicity values of the latter are equal to or less than the hydrophobicity values of DQβ49–56 in the hydrophobic alleles, i.e., groups 2 and 3 (Tables I and II).

The CD4-Binding Region

This region has been identified in part by site-directed mutagenesis and peptide binding as the sequence β134–148 on DR molecules (König *et al.*, 1992; Cammarota *et al.*, 1992) and shown by crystallography to reside on the surface of the β2 domain of DR1 (Brown *et al.*, 1993). It probably contains as well apposed residues from the α chain of the second DR molecule in the homodimer (Brown *et al.*, 1993), as corroborated by site-directed mutagenesis in selected residues from the mouse MHC class II α2 domain (König *et al.*, 1995). Indeed, in the homodimer of heterodimers residues from $(\beta 2)_A$ and $(\alpha 2)_B$ on one side and from $(\alpha 2)_A$ and $(\beta 2)_B$ on the diametrically opposite side form two identical grooves per homodimer of DR or DQ molecules (Figs. 1 and 8). The corresponding sequence in DQB molecules has three or four nonconservative substitutions: 135Gly → Asp, 139Lys → Thr, 140Ala → Thr (in 6/12 alleles with known sequences in this region, although this same substitution also occurs in about half of DR alleles as well (Marsh and Bodmer, 1995)), and 146Gly → Pro (Fig. 3). The modeled loop of this region in the DQ molecule shows that these substitutions do not alter the overall structure as defined in the DR1 molecule (Fig. 7). There are only two nonconservative substitutions in the apposed α chain residues participating in the formation of the CD4-binding region: 130 Pro → Ser and 175Glu → Gln or Lys in some alleles (numbering in DQα terms), while other DQ alleles retain the glutamate residue.

In a study just published, Fleury *et al.* (1995) have documented the involvement of the β-sheet segment DRβ180–189 in the binding of the CD4 molecule to DR, by pointing to dramatic differences in binding, arising from polymorphisms of various DR alleles in this region. Indeed this stretch is in β-pleated sheet conformation in DR and modeled DQ molecules (Fig.



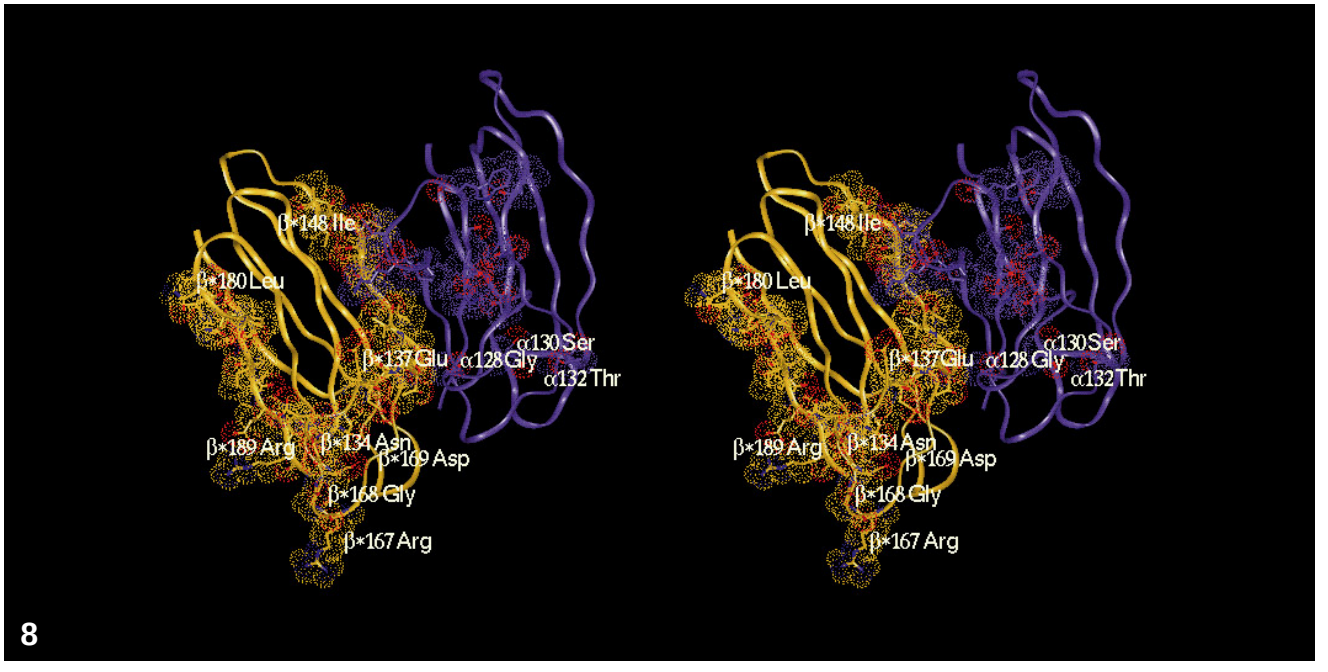


FIG. 6. Stereo view of the α2β2 dimerization domain in the DQ homodimer. Shown on the left of the figure is the backbone of the β2 domain of one DQ molecule (orange), on the top part a section (α80–120) of the α2 domain of this same molecule (red), and on the right the α2 domain of another DQ molecule (purple). To indicate that this structure originates from two heterodimers in a homodimer, the residues of the heterodimer in the left are starred (α* or β*). The crucial residues participating in the formatiuon of the homodimer are shown as van der Waals surfaces, with the carbon atoms having the colors of the respective chains from which they originate. In these residues, oxygen atoms are depicted in red and nitrogen atoms are depicted in deep blue. For clarification purposes not all residues participating in these interactions are labeled.

FIG. 8. Stereo view of the CD4-binding region of the DQ homodimer and the RGD loop. The orientation is identical to that of Fig. 6 and the picture concerns the identical backbone of the molecules. Color and numbering conventions are identical to those for Fig. 6. The crucial amino acids considered to be participating in CD4 recognition (Fig. 3) are shown as van der Waal surfaces. The CD4-binding stretches of β*134Asn to β*148Ile and β*180 Leu to β*189 Arg are demarkated by their first and last residues. The residue β*137Glu considered crucial for CD4 binding is also identified. Also identified are three residues from the apposed α2 domain of the other heterodimer, shown by site-directed mutagenesis to be crucial for CD4 binding by MHC class II molecules (König *et al.*, 1995). The three residues in the β*167–169 RGD loop are identified individually.

TABLE II
Hydrophobic Sequences on the Exterior of Oligomeric Water-Soluble Proteins

Protein	Region	Sequence	Hydrophobicity kcal/mole
Hemoglobin Sβ	Hydrophobic patch	3–6 LTPV	+6.4
Hemoglobin Aβ	Counterpart	3–6 LTPE	−4.4
Retinoid X[a] receptor α	Heptad repeat No. 9	418–425 LLLRLPAL	3.1
NF-κB p50[b]	Homodimerization region	267Y, 310V, 269L, 251V 304H, 308A, 307F	9.6
Serum response element[c]	Homodimerization region	180–188 TQVLLLVA	12.3
TNFα/β[d]	Trimerization region	12–18 PVAHVVA	7.2
		29–35 PAAHLIG	6.9
		55–64 LYLIYSQVLF	12.9
		71–80 IYFVYSQVVF	13.4
		117–124 PIYLGGVF	13.3
		131–138 SMYHGAAF	7.2
		144–157 FAESGQVYFGIIAL	11.6
		159–172 LSPST-VFFGAFAL	25.7
Tick-borne virus glycoproteins[e]	Homodimerisation region	257–268 LGDQTGVLLKAL	−5.5
	Interdomain contact	99–108 RGWGNH(G)CGLF	−6.7 (−2.7)
Human serum amyloid P[f]	Pentamerisation regions	111–118 TPLVKKGL	−7.0
		194–203 YVIIKPLVWV	9.0

[a] Bourguet *et al.* (1995).
[b] Ghosh *et al.* (1995a).
[c] Pellegrini *et al.* (1995).
[d] Eck and Sprang (1989) and Banner *et al.* (1993). In each homologous sequence pair the top line refers to TNFα sequence and the bottom line to TNFβ sequence. −, no amino acid in corresponding position.
[e] Rey *et al.* (1995).
[f] Emsley *et al.* (1994).

8), in a direction that is antiparallel to that of β140–148. All 13 DQB alleles with known sequences in this region differ from the vast majority of DR alleles in several residues: β180 (Leu → Val), β181 (Thr → Gln), β182 (in 6/12 there is a Ser → Asn substitution), β184 (Leu → Ile), and β185 (in 4/12 there is a Thr → Ile substitution) (Figs. 3 and 8). The substitutions in the various DQ alleles in the stretch of β180–189 do not invovle the same residues as in the case of the nonbinding of DR alleles in the study of Fleury *et al.* (1995). They do nonetheless affect some of the same positions (e.g., β180 and β181).

The β167–169 RGD Loop

Ten of thirteen DQB alleles with known sequences in the β2 domain have an RGD loop in position β167–169 (Figs. 1, 3, and 8). Such loops are prominent in many proteins of the cell membrane or the extracellular matrix that participate in cell adhesion interactions (integrins, etc; see Hynes, 1992, for a review). The orientation of the arginine and aspartate residues in the RGD loop of modeled DQ molecules is such that the two residues are part from each other and not in a position to directly form a salt bridge (Fig. 8). The other three alleles with known sequences in the β2 domain have a β167Arg → His substitution, yielding a sequence with no known cell adhesion function. This RGD loop is located about 10 amino acid residues above the plane of the plasma membrane of the cell bearing the DQ molecule. A search through the EMBL data bank has shown that 404 membrane proteins (integral or peripheral) from all species have RGD sequences, and of these 42 integral membrane proteins contain one or more RGD tripeptides in their extracellular domain. Many of the latter proteins so identified have a well-known role in cell adhesion via interactions of such RGD sequences with integrins or other proteins of the cell membrane and extracellular matrix (e.g., fibronectin and vitronectin). A considerable fraction of these integral membrane proteins are bona fide receptors (Ouzounis, Eliopoulos, and Papadopoulos, in preparation). We did not identify any other cell adhesion peptide sequences (e.g., DGEA, or EILDV, or GPRP, or KQAGDV, according to Hynes, 1992) in any of the DQαβ allelic sequences.

DISCUSSION

The modeled structure of the human class II histocompatibility antigens HLA-DQ shows many similarities to the crystallographically determined structure of its homologue DR1, yet with several distinct differences. These differences are propably of fucntional importance. The antigen-binding groove

has nearly the same strategically positioned polar residues on the two α-helices oriented in a way to form hydrogen bonds with amide and carbonyl groups of the trapped antigenic peptide. Thus the nonspecific manner of binding to the antigenic peptide backbone exhibited by HLA-DR molecules is maintained in all HLA-DQ molecules known. The specificity in antigenic peptide binding is provided by the various pockets that are formed within the groove. Our modeling shows that the physicochemical character of the groove in DQ molecules is different from that of DR, accommodating thus a different peptide motif. In particular, we find that the most prominent pockets in the groove are the second and the fifth. The second pocket accommodates a variety of residues, mostly hydrophobic ones because of its hydrophobic floor; yet, depending on the amino acids at positions β70, β71, and β74 (all positions polymorphic in DQB) that line the "mouth" of this pocket, different residues might be accommodated. The fifth pocket is as large as if not larger in DQ than its counterpart in DR1, since position β9 of the former is occupied by Tyr, Phe, or Leu, compared to Trp for the latter (Fig. 5C). Occasionally the fourth pocket also appears prominent, yet smaller in size than the previous two. By contrast, the first pocket in DQ molecules is not large enough to be an anchor, as it is in DR molecules. Furthermore, it is dimorphic, being either amphiphilic or hydrophilic with modifications in each case depending on the residue occupying position α34. The peptide sequences that would be expected to bind the DQ molecules would have a bulky hydrophobic residue (anchor(s)) in relative position 3 or 4 and/or 9 and a hydrophilic or amphiphilic residue in relative position 1. These conditions need not all be satisfied in one peptide. As already shown in the case of DR1 and the HA306–318 antigenic peptide, the very tight fit of 308Tyr in the first pocket of DR1, overcomes the energetically unfavorable binding of some of the other peptide residues in the remaining pockets of DR1.

Results consistent with the physicochemical character of the DQ-antigen-binding groove that we propose from our models have been obtained by pool sequencing of peptides eluted from DQ7: there was a clustering of hydrophilic residues in peptide relative position -2 (corresponding to our position 1) and anchoring with bulky hydrophobic residues in their position 1 (corresponding to our position 3) (Falk *et al.*, 1994). The other secondary anchors were relative positions 5 and 7 (corresponding to our positions 7 and 9, respectively). In addition, the individual sequencing of six intact peptides isolated from the groove of DQ7 shows that such peptides fit the basic pocket motif outlined here and have residues extending on either side of the groove. Furthermore, measuring the binding of selected peptides to purified DQ molecules as shown that the key residues from the antigenic peptide that determine binding are in the middle, corresponding to anchoring in the second pocket (Sidney *et al.*, 1994; Sinigaglia, 1995). Very recently binding studies to select DQ alleles showed that differences in only a few residues of the α or the β chain (the other chain being identical) could result in selective binding of three different antigenic peptides (Kwok *et al.*, 1995). The authors pointed out that most of the differing DQαβ residues are located in positions of presumed pockets (by analogy to DR1, e.g., β57); we in fact find nearly all such residues to occupy positions in pockets. The results of these studies have been confirmed by testing the binding of antigenic peptides from the putative diabetes autoantigen of glutamic acid decarboxylase, GAD (Baekkeskov *et al.*, 1990), to select DQ molecules involved in the susceptibility to or protection from type 1 diabetes (Kwok *et al.*, 1996a,b). It was found that peptide p34 (IARFKMFPEVKEK) from human GAD65 bound to DQA1*0301/B1*0302 (DQ3.2) but not to DQA1*0301/B1*0301 (DQ3.1). The two alleles differ in four positions only in the α1β1 domain, i.e., β13G → A, β26L → Y, β45G → E, and β57A → D, all except the third within the antigen-binding groove. By testing variant peptides substituting different residues at each position it was found that many single substitutions were allowed, except arginine at positions 1, 4, 6, and 9. Also, only Glu or Asp was allowed in position 9 (Kowk *et al.*, 1996a,b). However, a change of F → A in position 4 of the peptide maintained binding to DQ3.2, without any binding to DQ3.1 or DQ3.3. Furthermore, a change of 9E → A in the peptide allowed binding to DQ3.3 (DQA1*0301/B1*0303) and abolished binding to DQ3.1. The DQ3.3 allele differs from the DQ3.2 allele by the β57A → D substitution. Allowing both substitutions (4F → A and 9E → A) maintained binding to DQ3.3, allowed binding to DQ3.1, and abolished binding to DQ3.2. These results are explained only if the pockets are at peptide positions 1, 4, 6, and 9. Having acidic residues in β45(E) and β57(D), as in DQ3.3, dramatically changes the electrostatic potential of the DQ molecule (Sanjeevi *et al.*, 1995), thus not permitting the 9E of the peptide (negatively charged) to be accommodated in the fifth pocket; the change of 9E → A permits binding of this mutant peptide to DQ3.3 and potentially to DQ3.1. In DQ3.2, the change of β57 to A reduces the repulsive electrostatic effect, so 9E is accommodated well in the fifth pocket; also, 4F fits well in the second pocket. In DQ3.1 the β13A residue of the second pocket does not permit accommodation of 4F from the peptide, something allowed when this peptide residue is changed to A. These results are consistent with the pocket motifs

of DQ proposed here as a result of our modeling studies. They do show that in DQ binding there is no dominant pocket, but rather a concert of at least two such residue-binding pockets that may determine peptide binding. By contrast, all DR molecules (Jardetzky *et al.*, 1990; Falk *et al.*, 1994; Marshall *et al.*, 1994; Stern *et al.*, 1994; Stern, 1995; Ghosh *et al.*, 1995b) show preference for peptides where relative position 1 (anchor) is occupied by an aromatic residue, or an aliphatic residue, depending on the amino acid in position β86: glycine favors the former, while valine the latter (Verreck *et al.*, 1993). The other pockets of DR molecules do not seem to achieve the prominence of the first one (Stern *et al.*, 1994; Stern, 1995; Ghosh *et al.*, 1995b), although the fifth pocket can accommodate a bulky hydrophobic residue such as leucine (Stern *et al.*, 1994). Occasionally however, strong electrostatic considerations may give a pocket considerable prominence. For example, the autoimmune disease pemphigus vulgaris in Ashkenzy Jews is linked to the allele DRB1*0402, which contains two negatively charged residues, β70Asp and β71Glu, in its second pocket. The peptides from the desmoglein 3 autoantigen that specifically bind to this DR allele contain a lysine or an arginine in relative position 4 (second pocket). Thus, it is extremely attractive for the particular antigen peptides to dock a lysine or an argine at this position (Wucherpfennig *et al.*, 1995). Similarly, the molecule DQA1*0501/B1*0201 (DQ2), linked to coeliac disease in the Caucasian population, prefers an acidic residue at the second pocket, since positions β70 and 71 are occupied by Arg and Lys, respectively (van de Wal *et al.*, 1996).

The ability of DQ molecules to form homodimers of heterodimers is modulated by the presence of the polymorphic patch β49–56.The preponderance of *in vitro* experimental work points to the activation of T lymphocytes and MHC class II-bearing antigen-presenting cells as mediated by cognate receptor recognition and subsequent receptor dimerization (Janeway, 1992; Brown *et al.*, 1993; Racciopi *et al.*, 1993; König *et al.*, 1995). The two hydrophobic variants of the β49–56 patch would tend to form dimers spontaneously by analogy to the situation with hemoglobin S, where the hydrophobic stretch β3–6 is well matched with the exposed hydrophobic stretch in the α chain of the deoxy state of the molecule (Noguchi and Schechter, 1985). On a per residue basis, this β49–56 patch in the two hydrophobic variants of DQ alleles is a hydrophobic as the intramembranous parts of integral membrane proteins (1.1–1.8 kcal/mole/residue for 22-residue stretches, Stryer, 1995). Our search of water-soluble proteins, or fragments thereof, with known three-dimensional structure has revealed few 8-residue stretches of comparable hydrophobicity exposed to the solvent (Table II). Such sequences were found in the dimerization stretches in the exterior of several water-soluble transcription factors (NF-κB p50 (Ghosh *et al.*, 1995a, and Muller *et al.*, 1995), human serum response ractor (Pellegrini *et al.*, 1995), retinoic acid-binding protein (Bourguet *et al.*, 1995)), the dimerization sections of tick-borne virus envelope glycoprotein (Ray *et al.*, 1995), the trimerization patches of TNFα and TNFβ (Eck and Sprang, 1989, and Banner *et al.*, 1993), and the pentamerization patches of human serum amyloid P (Emsley *et al.*, 1994)) (Table II). Thus, the hydrophobic DQβ49–56 patches would be expected to promote homodimer formation by virtue of their hydrophobicity. The amphiphilic and hydrophilic variants of this patch are positively charged, making the self-association into homodimers problematic. T lymphocyte receptor recognition could overcome the repulsion of arginines in the case of the amphiphilic variant because of the van der Waals interactions among the several leucine residues within the patch (Routsias and Papadopoulos, 1995). In the hydrophilic variant of the patch the T lymphocyte receptor might also perform a similar function, but here the tendency of the two patches to remain apart is greater because of the positive charges on β55Arg. The α2β2 homodimerization domain in all DQαβ alleles is nearly identical to that of DR1. Thus, in those alleles with hydrophobic β49–56 sequences, spontaneous dimerization via this patch could induce a like process in the α2β2 domain; dimerization induced by the T lymphocyte receptor would be well promoted in these alleles. By contrast, the remaining DQ alleles would probably have homodimerization in both regions induced with some difficulty by T lymphocyte receptor, with those having arginine in position β55 showing the greatest difficulty to do so.

There are no polymorphisms in the residues of the α and the β chains responsible for the α2β2 interheterodimer contacts (Figs. 3 and 4), except for position DQα160, where one allele (A1*0302) has an aspartate and another (A1*0503) a serine, instead of alanine as in the other seven alleles with known sequences in the α2 domain. Furthermore, in comparison to DRαβ, of the 15 residues maintaining these interheterodimer contacts there is only one conservative substitution (DQα161E → D), one semiconservative (DQα114K → Q), and a nonconservative one (DQα160T → A). The glutamate residue in position DRα88 (DQα91) has been shown by site-directed mutagenesis as essential for activation of T lymphocyte clones, presumed to be via dimerization (Goodman *et al.*, 1995). The mutated DRa91Glu → Ala allele can still bind antigen and form an αβ heterodimer; thus, the defect in T lymphocyte activa-

tion probably rests in the inability of such a mutant to form a homodimer of heterodimers. By contrast, a mutation in DRα111Lys → Ala or Glu does not abrogate the ability of the DR molecule to activate T lymphocyte clones. This is indirect evidence that the semiconservative substitution at the equivalent position in all DQα alleles compared to DRα (Lys → Gln) does not affect the ability of DQαβ to form homodimers.

There are three or four amino acid substitutions in the first identified CD4-binding stretch of β134–148 in DQ alleles, in comparison to DR1. The substitutions are in positions not shown to be crucial for CD4 binding by site-directed mutagenesis. The latter showed Glu137 and Val143 to be of the essence, while a threonine to alanine substitution at position 140 showed no change in affinity for CD4 (König *et al.,* 1992). It appears from the available evidence that CD4 binding to MHC class II molecules is promoted by their homodimerizatiuon, and the residues on DQα apposed to the β134–148 stretch are probably of importance in this process (König *et al.,* 1995, and Fig. 8). The differences between the various DQA alleles in these positions and the corresponding residues from DRα are very few (Fig. 2). The recent finding that the peptide stretch from DRβ180 to 189, especially the residues in positions β180 and 181, contributes significantly to DR binding to CD4 raises the possibility of differences in binding to CD4 by DR compared to DQ alleles, as there are several substitutions in this region in the DQ alleles (Fig. 3). These substitutions are not conservative, and in some DQ alleles they encompass 5/10 residues in this stretch. Only binding experiments will determine the effect of these replacements on CD4 ligation and subsequent T lymphocyte and antigen-presenting cell activation events. In another set of experiments Huard *et al.* (1995) have shown that the interaction of soluble human CD4 with HLA-DR molecules on Daudi (B-lymphoid-like) cells has an affinity of 3 μ*M*, while the affinity of lymphocyte activation gene (LAG)-3, also in soluble form, for HLA-DR is much higher, ca. 60 n*M*. The latter gene product was shown by these authors to compete with CD4 for the interaction with HLA-DR molecules. However, no such competition was observed on CD4-restricted T lymphocyte clones that were specific for HLA-DR plus peptide antigen. These results are consistent with the interpretation that in intact cells the ligation of T lymphocyte receptor with HLA-DR renders the interaction of the latter, and by extension of HLA-DQ, with CD4, immune from competition with LAG-3.

The necessity of HLA-DR homodimerization for cognate T lymphocyte activation was further documented by showing that stable binding of CD4 to MHC class II molecules is observed when the former oligomerize via their membrane proximal extracellular domains D3 and D4 (Sakihama *et al.,* 1995). The biological importance of the homodimerization of MHC class II αβ heterodimers had been established in part from earlier studies, where cross-linking of these molecules on B lymphocytes with specific antibodies led to the activation of these cells, while no such effect was observed with the monovalent fragment of the same antibodies (Mooney *et al.,* 1989; Lane *et al.,* 1990; Koulova *et al.,* 1991). It seems that there is a variety of cell surface receptors (e.g., human growth hormone receptor (de Vos *et al.,* 1992), several growth factor receptors (Brugge, 1994), and cytokine receptors (Ihle *et al.,* 1994)) that upon ligand binding must undergo homodimerization in order for signal transduction to take place.

The β167–169 RGD loop exhibited by 10/13 DQβ alleles with known sequences in this region requires further investigation. Such loops have been shown to be of essence in the interaction of several cell adhesion molecules (vitronectin, fibronectin, tenascin (Leahy *et al.,* 1992) with their cell surface ligands called integrins (Hynes, 1992)), as well as natural inhibitors of this process (mainly snake venoms under the name of disintegrins (Saudek *et al.,* 1991; Adler *et al.,* 1992)). In all the cases where the three-dimensional structure of the molecule is known, the tripeptide was in a loop structure (Leahy *et al.,* 1992; Haas and Plow, 1994). Furthermore, linear RGD tripeptides do not inhibit this process as well as cyclical RGD peptides (Samanen *et al.,* 1991). When RGD loops were engineered at turns of the polypeptide chains of lysozyme (a protein with no function in cell adhesion), the mutein inhibited the interaction of fibrinogen with integrin $\alpha_{IIb}\beta_3$ at nanomolar concentrations (Yamada *et al.,* 1994). A recent comparison of the conformation of RGD loops from three different proteins involved in cell adhesion showed that in general the Asp and Arg residues are kept apart from each other (Grimes *et al.,* 1995), as is the case here with modeled HLA-DQ structures (Fig. 8). This sequence is also found in rat and swine MHC class II proteins (RT1.B and SLA-DQ, respectively) in equivalent positions to DQβ167–169 (Holowachuck and Greer, 1989; Gustafson *et al.,* 1990). Whether this RGD loop confers upon the DQ molecules that possess it the ability to participate in cell adhesion interactions is a matter for further investigation. While the loop occurs only 10 residues above the membrane plane of the cell bearing the DQ molecule, it is known that integrins, the primary RGD ligands, have an extracellular arm that is nearly 200 Å long (Hynes, 1992). This is greater than the distance between the antigen-presenting cell and a T lymphocyte, when these two cells make contact

via the MHC–peptide complex and the T lymphocyte receptor molecule (where the extracellular portion of each of these molecules is nearly 75 Å long (Brown *et al.*, 1993; Fields *et al.*, 1995)).

The HLA-DQ molecules are expressed in distinct areas of the human fetal thymus and probably function as restriction elements for CD4+ T lymphocytes (Janossy *et al.*, 1986). Furthermore, the percentage of T lymphocytes in human peripheral blood that recognize antigen presented by HLA-DQ molecules is remarkably high (Fujisawa *et al.*, 1991). So the question remains of the mode of action of these class II histocompatibility proteins as they are implicated in such diverse processes as autoimmunity and immunosuppression. There are several implications of the modeling results for the DQ molecules concerning the immunological function of the latter. It is obvious from our studies that DQ molecules require a different peptide motif for binding from that of DR molecules. The spontaneity of homodimer formation in some DQ alleles probably signifies a fast response to given antigenic peptides, once the DQ molecule has been loaded with the peptides. By contrast, DQ alleles that have difficulty in forming homodimers probably do not activate cognate T lymphocyte receptor ligands as well. The RGD loop probably adds a cell adhesion function to many DQ alleles, perhaps strengthening the interaction of antigen-presenting cells with cognate T lymphocytes or perhaps with other cells that might interact with DQ-bearing antigen-presenting cells (Routsias and Papadopoulos, 1995). Two studies have correlated well the surface electrostatic potential of the $\alpha 1\beta 1$ domain and polymorphic structural features of the extracellular part of the molecule to susceptibility or dominant resistance of HLA-DQ alleles to type 1 diabetes (Sanjeevi *et al.* (1995) and Routsias and Papadopoulos (1995). It should be added however, that the short cytoplasmic tails of DQα and β chains are very different from their counterparts in DRα and β (Marsh and Bodmer, 1995), signifying perhaps different mechanisms of signal transduction once recognition by T lymphocyte antigenic receptor and dimerization is in place. Furthermore, allelic differences in the gene regulatory sequences of DQA and DQB genes have been recorded (Andersen *et al.*, 1991; Morzycka-Wroblewska *et al.*, 1993; Woolfrey and Nepom, 1995). The level of expression of DQ alleles could very well influence their functional capacity in a given antigenic response. The concert of these structural subtleties probably constitutes the essence of the functional differences between HLA-DQ molecules and other human class II molecules of the major histocompatibility complex.

We thank Professor Don C. Wiley for initially suppying the unrefined coordinates of DR1 and later the refined coordinates of the complex of DR1 and HA306-318 prior to publication, Dr. Doug Unlien for help in locating DQ sequences, the Division of Organic Chemistry and Biochemistry of the University of Ioannina for allowing us to use their Silicon Graphics Iris workstation, Ms. Paraskevi Mitlianga for help in drawing Fig. 7, Ms. Katherine Stefopoulou for drawing Fig. 9 of the unrevised manuscript, Professors Å Lernmark, G. Panayi, and E. Thorsby and Dr. J. Lanchbury for helpful suggestions, and Dr. Chris Bartsocas for encouragement, Also, thanks are due to the many workers (too numerous to attribute individually in the references) who cloned and sequenced the several DQA and DQB alleles used in our modeling and the subsequent analysis. This work was supported in part by a grant from the EU to G.K.P. under the Biotech program.

REFERENCES

Adler, M., Lazarus, R. A., Dennis, M. S., and Wagner, G. (1991) Solution structure of kistrin, a potent platelet aggregation inhibitor and GP IIb-IIIa antagonist, *Science* **253,** 445–448.

Altman, D. M., Sansom, D., and Marsh, S. G. E. (1991) What is the basis of HLA-DQ associations with autoimmune disease? *Immunol. Today* **12,** 267–270.

Andersen, L. C., Beaty, J. S., Nettles, J. W., Seyfried, C. E., Nepom, G. T., and Nepom, B. S. (1991) Allelic polymorphism in transcriptional regulatory regions of HLA-DQB genes, *J Exp. Med*. **215,** 181–192.

Baekkeskov, S., Aanstood, H.-J., Christgau, S., Reetz, A., Solimena, M., Cascalho, M., Folli, H., Richter-Olesen, H., and de Camilli, P. (1990) Identification of the 64K autoantigen in insulin-dependent diabetes as the GABA-synthesizing enzyme glutamic acid decarboxylase, *Nature* **345,** 151–156.

Bairoch, A., and Boeckman, B. (1991) The SWISS-PROT protein sequence data bank, *Nucleic Acids Res.* **19,** 2247–2249.

Banner, D. W., D'Arcy, A., Janes, W., Gentz, R., Schoenfelt, H.-J., Broger, C., Loetscher, H., and Lesslauer, W. (1993) Crystal structure of the soluble human 55 kd TNF receptor-human TNFβ complex: implications for TNF receptor activation, *Cell* **73,** 431–445.

Bjorkman, P. J., Saper, M. A., Samroui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C. (1987) Structure of the human class I histocompatibility antigen HLA-A2, *Nature* **329,** 506–512.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995) Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-α, *Nature* **375,** 377–382.

Briata, P., Radka, S. F., and Sartoris, S. (1989) Alternative splicing of HLA-DQ transcripts and secretion of HLA-DQ β chain protein: Allelic polymorphism in splicing and polyadenylation sites, *Proc. Natl. Acad. Sci. USA* **86,** 1003–1007.

Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J., and Wiley, D. C. (1988) A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules, *Nature* **332,** 845–850.

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1, *Nature* **364,** 33–39.

Brugge, J. S. (1994) Protein:protein interactions involved in intracellular signal transduction, *Chem. Biol.*, **introductory issue,** 15 April, xii–xiii.

Cammarota, G., Scheirle, A., Tacaks, B., Doran, D. M., Knorr, R., Bannworth, W., Guardiola, J., and Sinigaglia, F. (1992) Identification of a CD4 binding site on the $\beta 2$ domain of HLA-DR molecules, *Nature* **356,** 799–801.

Campbell, R. D., and Trowsdale, J. (1993) Map of the human MHC, *Immunol. Today* **14,** 349–352.

de Vos, A. M., Ultsch, M., and Kossiakof, A. A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex, *Science* **255,** 306–312.

Eck, M. J., and Sprang, S. R. (1989) The structure of tumor necrosis factor-α at 2.6 Å resolution: implications for receptor binding, *J. Biol. Chem.* **264,** 17595–17605.

Emsley, J., White, H. E., O'Hara, B. P., Oliva, G., Srinivasan, N., Tickle, I. J., Blundell, T. L., Pepys, M. B., and Wood, S. P. (1994) Structure of pentameric serum amyloid P component, *Nature* **367,** 338–345.

Engelman, D. M., Steitz, T. A., and Goldman, A. (1986). Identifying non-polar transbilayer helices in amino acid sequences of membrane proteins, *Annu. Rev. Biophys. Biophys. Chem.* **15,** 321–353.

Falk, K., Rötzschke, O., Stefanovic, S., Jung, G., and Rammensee, H.-G. (1994) Pool sequencing of natural HLA-DR, DQ and DP ligands reveals detailed peptide motifs, constraints of processing and general rules, *Immunogenetics* **39,** 230–242.

Fields, B. A., Ober, B., Malchiodi, E. L., Lebedeva, M. I., Braden, B. C., Ysern, X., Kim, J.-K., Shao, X., Ward, E. S., and Mariuzza, R. A. (1995) Crystal structure of the V_α domain of a T cell antigen receptor, *Science* **270,** 1821–1824.

Fleury, S., Thibodeau, J., Croteau, G., Labreque, N., Aronson, H.-E., Cantin, C., Long, E. O., and Sêkaly, R.-P. (1995) HLA-DR polymorphism affects the interaction with CD4, *J. Exp. Med.* **182,** 733–741.

Fossum, B., Gedde-Dahl III, T., Hansen, T., Eriksen, J. A., Thorsby, E., and Gaudernack, G. (1993) Overlapping epitopes encompassing a point mutation (12Gly → Arg) in p21 ras can be recognised by HLA-DR, -DP, and -DQ restricted T lymphocytes, *Eur. J. Immunol.* **23,** 2687–2691.

Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1992) Crystal structures of two viral peptides in complex with murine MHC Class I H-$2K^b$, *Science* **257,** 919–927.

Fremont, D. H., Stura, E. A., Matsumura, M., Peterson, P. A., and Wilson, I. A. (1995). Crystal structure of an H-$2K^b$-ovalbumin peptide complex reveals the interplay of primary and secondary anchor positions in the major histocompatibility complex binding groove. *Proc. Natl. Acad. Sci. USA* **92,** 2479–2483.

Fremont, D. H., Hendrickson, W. A., Marrack, P., and Kappler, J. (1996) Structures of an MHC class II molecule with covalently bound single peptides, *Science* **272,** 1001–1004.

Fujisawa, K., Kamikawaji, N., Yasunami, M., Kimura, A., Nishimura, Y., and Sasazuki, T. (1991) High precursor frequency of human T lymphocytes reactive to HLA-DQ molecules expressed on mouse L cells transfectants, *Eur. J. Immunol.* **21,** 2341–2347.

Gedde-Dahl III, T., Fossum, B., Eriksen, J. A., Thorsby, E., and Gaudernack, G. (1993) T lymphocyte clones specific for p21 ras-derived peptides: characterization of their fine specificity and HLA restrictions, *Eur. J. Immunol.* **23,** 754–760.

Germain, R. N. (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for lymphocyte activation, *Cell* **76,** 287–299.

Ghosh, G., Van Duyne, G., Ghosh, S., and Sigler, P. B. (1995a) Structure of NF-κB p50 homodimer bound to a κB site, *Nature* **373,** 303–310.

Ghosh, P., Amaya, M., Mellins, E., and Wiley, D. C. (1995b) The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3, *Nature,* **378,** 457–462.

Goodman, S., Sawada, T., Barbosa, J. A., Cole, B., Pergolizzi, R., Silver, J., Mellins, E., and Chang, M.-der Y. (1995) Mutational analysis of two DRα residues involved in dimers of HLA-DR molecules, *J. Immunol.* **155,** 1210–1217.

Grimes, J., Basak, A. K., Roy, P., and Stuart, D. (1995) The crystal structure of bluetongue virus VP7, *Nature* **373,** 167–170.

Gustafsson, K., LeGuern, C., Hirsch, F., Germana, S., Pratt, K., and Sachs, D. H. (1990) Class II MHC genes of miniature swine IV: Characterization and expression of two allelic class II DQB cDNA clones, *J. Immunol.* **145,** 1946–1951.

Haas, T. A., and Plow, E. F. (1994) Integrin-ligand interactions: a year in review. *Curr. Opin. Cell Biol.* **6,** 656–662.

Hirayama, K., Matsushita, S., Kikuchi, I., Iuchi, M., Ohta, N., and Sasazuki, T. (1987) HLA-DQ is epistatic to HLA-DR in controlling the immune response to schistosomal antigens in humans, *Nature* **327,** 426–429.

Holowachuk, E. W., and Greer, M. K. (1989) Unaltered class II histocompatibility antigens and pathogenesis of IDDM in BB rats, *Diabetes* **38,** 267–271.

Huard, B., Prigent, P., Tournier, M., Bruniquel, D., and Triebel, F. (1995) CD4/major histocompatibility complex class II interaction analyzed with CD4- and lymphocyte activation gene-3 (LAG-3)-Ig fusion proteins, *Eur. J. Immunol.* **25,** 2718–2721.

Hynes, R. O. (1992) Integrins: Versality, modulation and signaling in cell adhesion, *Cell* **69,** 11–25.

Ihle, J. N., Witthuhn, B. A., Quelle, F. W., Yamamoto, K., Thierfelder, W. E., Kreider, B., and Silvennoinen, O. (1994) Signaling by the cytokine receptor superfamily: JAKs and STATs, *Trends Bioch. Sci.* **19,** 222–227.

Janeway, C. A., Jr. (1992) The T cell receptor as a multicomponent signalling machine: CD4/CD8 coreceptors and CD45 in T cell activation, *Annu. Rev. Immunol.* **10,** 645–674.

Janossy, G., Bofill, M., Poulter, L. W., Rawlings, E., Burford, G., Navarrette, C., Ziegler, A., and Kelemen, E. (1986) Separate ontogeny of two macrophage-like accessory cell populations in the human fetus, *J. Immunol.* **136,** 4354–4361.

Jardetzky, T. S., Gorga, J. C., Busch, J., Rothbard, J., Strominger, J. L. and Wiley, D. C. (1990) Peptide binding to DR1: A peptide with most residues substituted to alanine retains binding, *EMBO J.* **9,** 1797–1803.

Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Chi, U.-i., Stauffacher, C., Strominger, J. L., and Wiley, D. C. (1994) Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. *Nature* **368,** 711–718.

Johansen, B. H., Buus, S., Vartal, F., Viken, H., Eriksen, J. A., Thorsby, E., and Sollid, L. M. (1994) Binding peptides to HLA-DQ molecules: peptide binding properties of the disease-associated HLA-DQ (α1*0501, β1*0201) molecule, *Int. Immunol.* **6,** 453–461.

Jones, C. M., Lake, R. A., Lamb, J. R., and Faith, A. (1994) Degeneracy of T lymphocyte receptor recognition of an influenza virus hemagglutinin epitope restricted by HLA-DQ and -DR class II molecules, *Eur. J. Immunol.* **24,** 1137–1142.

Kabat, E. A., Wu, T. T., Reid-Miller, M., and Gottesman, K. S. (1991) Sequences of Proteins of Immunological Interest, 5th ed., Public Health Service, NIH, Washington, DC.

Kim, J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) Toxic shock syndrome toxin-1 complexed with a class II major histocompatibility molecule HLA-DR1, *Science* **266,** 1870–1874.

König, R., Huang, Y.-L. and Germain, R. N. (1992) MHC class II interaction with CD4 mediated by a region analogous to the MHC class I binding site for CD8, *Nature* **356,** 796–798.

König, R., Shen, X., and Germain, R. N. (1995) Involvement of both major histocompatibility complex class II α and β chains in

CD4 function indicates a role for ordered oligomerisation in T lymphocyte activation, *J. Exp. Med.* **182,** 779–787.

Koulova, L., Clark, E. A., Shu, G., and Dupont, B. (1991) The CD28 ligand B7/BB1 provides costimulatory signal for alloactivation of $CD4^+$ T lymphocytes, *J. Exp. Med.* **173,** 759–762.

Kwok, W. W., Nepom, G. T., and Raymond, F. C. (1995) HLA-DQ polymosphisms are highly selective for peptide binding interactions, *J. Immunol.* **155,** 2468–2476.

Kwok, W. W., Domeier, M. E., Raymond, F. C., Byers, P., and Nepom, G. T. (1996a) Allele-specific motifs characterize HLA-DQ interactions with a diabetes associated peptide derived from glutamic acid decarboxylase, *J. Immunol.* **156,** 2171–2177.

Kwok, W. W., Domeier, M. E., Johnson, M. L., and Nepom, G. T. (1996b) HLA-DQB1 codon 57 is critical for peptide binding and recognition, *J. Exp. Med.* **183,** 1253–1258.

Lane, P. J. L., McConnell, F. M., Schieven, G. L., Clark, E. A., and Ledbetter, J. A. (1990) The role of class II molecules in human B cell activation: Association with phosphatidyl inositol turnover, protein tyrosine phosphorylation and proliferation, *J. Immunol.* **144,** 3684–3692.

Leahey, D. J., Hendrickson, W. A., Aukhil, I., and Erickson, H. P. (1992) Structure of a fibronectin type III domain for tenascin phased by MAD analysis of the selenomethionyl protein, *Science* **258,** 987–991.

Lotteau, V., Teyton, R., Borroughs, D., Charron, D. (1987) A novel HLA class II molecule (DRαDQβ) created by mismatched isotype pairing, *Nature* **329,** 339–342.

Male, D., Champion, B., Cooke, A., and Owen, M. (1991) Advanced Immunology, 2nd ed., Gower, London.

Marsh, S. G. E., and Bodmer, J. G. (1995) HLA class II region nucleotide sequences, 1995, *Tissue Antigens* **46,** 258–280.

Marshall, K. W., Liu, A. F., Canales, J., Perahia, B., Jorgensen, B., Gantzos, R., Aguilar, B., Devaux, B., and Rothbard, J. (1994) Role of polymorphic residues in HLA-DR molecules in allele-specific binding of peptide antigens, *J. Immunol.* **152,** 4946–4957.

Matsumura, M., Fremont, D. H., Peterson, P. A., Wilson, I. A. (1992) Emerging principles for the recognition of peptide antigens by MHC class I molecules, *Science* **257,** 927–934.

Mooney, NB., Grillot-Courvalin, C., Hivroz, C., and Charron, D. (1989) A role for MHC class II antigens in B-cell activation, *J. Autoimmun.* **2**(Suppl.), 215–223.

Morzycka-Wroblewska, E., Harwood, J. I., Smith, J. R., and Kagnoff, M. S. (1993) Structure and evolution of the promoter regions of the DQA genes, *Immunogenetics* **37,** 364–372.

Muller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995) Structure of the NF-κB p50 homodimer bound to DNA, *Nature,* **373,** 311–317.

Noguchi, C. T., and Schechter, E. M. (1985) Sickle cell hemoglobin polymerisation in solution and in cells, *Annu. Rev. Biophys. Biophys. Chem.* **14,** 239–263.

Pellegrini, L., Tam, S., and Richmond, T. J. (1995) Structure of serum response factor core bound to DNA, *Nature* **376,** 490–498.

Racciopi, L., Ronchese, F., Matis, L. A., and Germain, R. N. (1993). Peptide-major histocompatibility complex class II complexes with mixed agonist/antagonist properties provide evidence for ligand-related differences in T-cell receptor-dependent intracellular signalling, *J. Exp. Med.* **177,** 1047–1060.

Ray, F. A., Heinz, F. X., Mandl, C., Kunz, C., and Harrison, S. C. (1995) The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution, *Nature* **375,** 291–298.

Routsias, J., and Papadopoulos, G. K. (1995) Polymorphic structural features of modelled HLA-DQ molecules segregate according to susceptibility or resistance to IDDM, *Diabetologia* **38,** 1251–1261.

Sakihama, T., Smolyar, A., and Reinherz, E. L. (1995) Oligomerisation of CD4 is required for stable binding to class II major histocompatibility complex proteins but not for interaction with human immunodeficiency virus gp120, *Proc. Natl. Acad. Sci. USA* **92,** 6444–6448.

Salgame, P., Convit, J., and Bloom, B. R. (1991) Immunological suppression by human $CD8^+$ T lymphocytes is receptor dependent and HLA-DQ restricted, *Proc. Natl. Acad. Sci. USA* **88,** 2598–2602.

Sanjeevi, C. B., Lybrand, T. P., DeWeese, C., Landin-Olsson, M., Kockum, I., Dahlquist, G., Sundqvist, G., Stenger, D., Lernmark, Å., and the members of the Swedish Childhood Diabetes Study (1995) Polymorphic amino acid variations in HLA-DQ are associated with systemic physical property changes and occurrence of IDDM, *Diabetes* **44,** 125–131.

Saudek, V., Atkinson, R. A. and Pelton, J. T. (1991) Three-dimensional structure of echistatin, the smallest active RGD protein, *Biochemistry* **30,** 7369–7372.

Sette, A., Adorini, L., Appella, E., Colon, S. M., Miles, C., Tanaka, S., Ehrhardt, C., Doria, G., Nagy, Z. A., Buus, S., and Grey, H. M. (1989) Structural requirements for the interaction between peptide antigens and $I\text{-}E^d$ molecules, *J. Immunol.* **143,** 3289–3294.

Sidney, J., Oseroff, C., del Guercio, M.-F., Southwood, S., Kroeger, J. I., Ishioka, G. Y., Sagaguchi, K., Appella, E., and Sette, A. (1994) Definition of a DQ3.1-specific binding motif, *J. Immunol.* **152,** 4516–4525.

Samanen, J., Ali, F., Romoff, T., Calvo, R., Sorenson, E., Vasko, J., Storer, B., Berry, D., Bennet, D., Strohsacker, M., Powers, D., Stadel, J., and Nicholls, A. (1991) Development of a small RGD peptide fibrinogen receptor antagonist with potent antiaggregatory activity in vitro, *J. Med. Chem.* **35,** 3114–3125.

Sinigaglia, F. (1995) Peptide binding by disease-associated MHC—class II molecules, Symposium (S6) lecture, 9th International Congress of Immunology, 23–29 July, San Francisco, CA.

Srinivasan, N., and Blundell, T. L. (1993) An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure, *Protein Eng.* **6,** 501–510.

St. Charles, R., Harrison, R. W., Bell, G. I., Pilkis, S. J., and Weber, I. (1994) Molecular model of human β-cell glucokinase built by analogy to the crystal structure of yeast hexokinase B, *Diabetes,* **43,** 784–791.

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1994) Crystal structure of a human class II MHC protein HLA-DR1 complexed with an influenza virus peptide, *Nature* **368,** 215–221.

Stern, L. J. (1995) A complex of HLA-DR1 with a peptide from HLA-A8, Symposium (S6) lecture, 9th International Congress of Immunology, 23–29 July, San Francisco, CA, USA, 23–29 July.

Stryer, L. (1995) Biochemistry, 4th ed., p. 284, Freeman, New York.

Trowsdale, J., and Campbell, R. D. (1992) Complexity in the major histocompatibility complex, *Eur. J. Immunogenet.* **19,** 43–55.

van de Wal, Y., Kooy, Y. M. C., Drijfhout, J. W., Amons, R., and Koning, F. (1996) Peptide-binding characteristics of the coeliac disease-associated DQ(α1*0501, β1*0201) molecule. *Immunogenetics*, in press.

Verreck, F. A. W., Termijtelen, A., and Koning, F. (1993) HLA-DRβ chain residue 86 controls DRαβ dimer stability, *Eur. J. Immunol.* **23,** 1346–1350.

Weber, I., Miller, M., Jaskolski, M., Leis, J., Skalka, A. M., and Wlodawer, A. (1989) Molecular modeling of HIV-1 protease and its substrate binding site, *Science* **243,** 928–931.

Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) Conserved folding in retrovial proteases: Crystal structure of a synthetic HIV-1 protease, *Science* **245,** 616–621.

Woolfrey, A. E., and Nepom, G. T. (1995) Differential transcriptional elements direct expression of HLA-DQ genes, *Clin. Immunol. Immunopathol.* **74,** 119–126.

Wucherpfennig, K. W., Yu, B., Bhol, K., Monos, D. S., Argyris, E., Karr, R. W., Ahmed, A. R., and Strominger, J. L. (1995) Structural basis for major-histocompatibility complex (MHC)-linked suscpetibility to autoimmunity: Charged residues of a single MHC binding pocket confer selective presentation of self peptides in pemphigus vulgaris, *Proc. Natl. Acad. Sci. USA* **92,** 11935–11939.

Yamada, T., Uyeda, A., Kidera, A., and Kikuchi, M. (1994) Functional analysis and modeling of a conformationally constrained Arg-Gly-Asp sequence inserted into human lysozyme, *Biochemistry* **33,** 11678–11683.