The class I and class II proteins of the human major histocompatibility complex

Jack L. Strominger[1,2] *& Don C. Wiley*[1,3]

Transplantation antigens, now called major histocompatibility complex (MHC) antigens or MHC glycoproteins, were discovered by Peter Gorer in the late 1930s. Graft rejection results from their polymorphism. Later they were found to give rise after whole-blood transfusion or during pregnancy to alloantibodies that recognize human leukocyte antigens (HLA). This discovery led to tissue typing, now widely employed in the matching of donors and recipients of organ grafts. The polymorphism posed the fundamental biological problem of the nature of the selection force which led to its evolution, since it could not have been exchange of surgical grafts. A related polymorphism was soon described in immune response (Ir) genes and, moreover, transplantation antigens and Ir genes were mapped to the same genetic region, the MHC on human chromosome 6 or murine chromosome 17. Through the work of many investigators, including our colleagues Peter Doherty, Emil Unanue and Rolf Zinkernagel, these two classes of molecules are now known to play the key role in the presentation of foreign protein antigens (processed to peptides) to the immune system. The rejection of grafts is a byproduct of this fundamental role in immunity. The polymorphism that evolved permits presentation of a large variety of foreign peptides to the immune system.

Work on the biochemistry of these molecules began about 1970. Initially, a quantitative assay was developed based on the inhibition by soluble HLA antigen of cytotoxicity mediated by pregnancy alloantisera in the presence of complement, and was then used to purify these molecules to homogeneity. Two prominent peaks were separated at the last step in the purification, gel filtration, but the assay detected the material in only one of these two peaks. SDS gels showed

---

Reprinted with permission from: Strominger JL, Wiley DC. The class I and II proteins of the human major histocompatibility complex. *JAMA* 1995; 274: 1074 1076. Copyright 1995, American Medical Association.

[1]Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, [2]Dana-Farber Cancer Institute, Division of Tumor Virology, 44 Binney Street, Boston, MA 02115, [3]Laboratory of Molecular Medicine, Department of Medicine, Howard Hughes Medical Institute, The Children's Hospital, 320 Longwood Avenue, Boston, MA 02115, USA.

HLA AND DISEASE THE MOLECULAR BASIS. Alfred Benzon Symposium 40, pages 38 45
Editors: Arne Svejgaard, Søren Buus, Lars Fugger, Munksgaard, Copenhagen 1997.

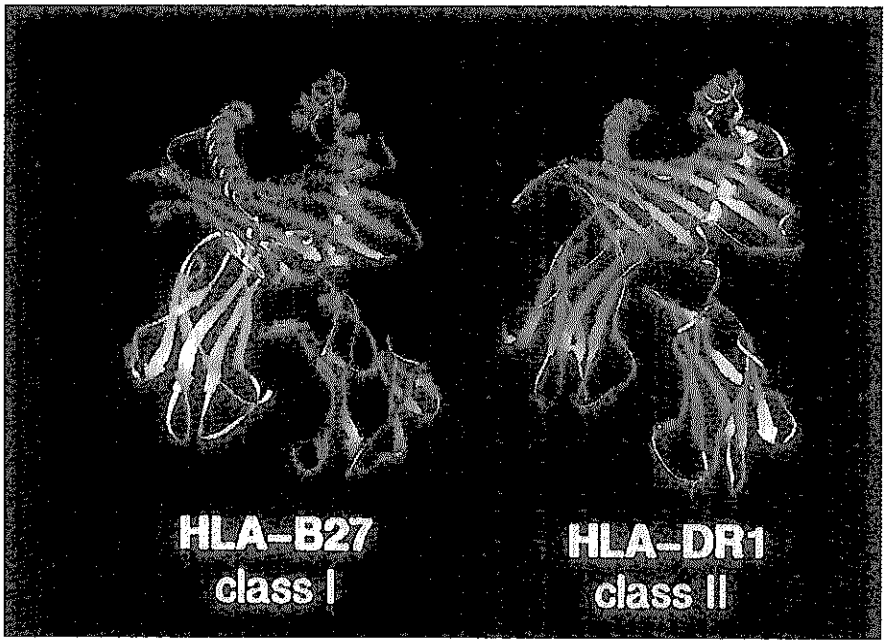

Fig. 1. Structures of class I and class II MHC proteins (2, 3, 8). Ribbon diagrams are shown. The clefts between the two helices at the top are the sites of peptide binding.

that each peak contained a two-chain heterodimer. Moreover, many properties suggested that the two molecules might be closely related. They are now called class I and class II MHC proteins, i.e. transplantation antigens and Ir gene products respectively. The first outline of the structures of class I and class II MHC proteins was developed from sequencing of the isolated proteins and of DNA that encodes these molecules. The two membrane proximal domains in each case were Ig-like domains; they were the first described members of what is now called the immunoglobulin superfamily of proteins. The two membrane distal domains were not Ig-like and carried the polymorphic determinants that characterize these proteins. Each molecule contained four external domains, but the connectivity of the domains was different in the two classes, three in the heavy chain and one in the light chain in class I, while there were two in each chain in class II (reviewed in ref. 1, 2).

STRUCTURES OF CLASS I AND II MHC PROTEINS

The 3-dimensional structure of a class I protein was determined in 1987 (3), and that of a class II protein in 1993 (4). As predicted, each has two membrane proximal

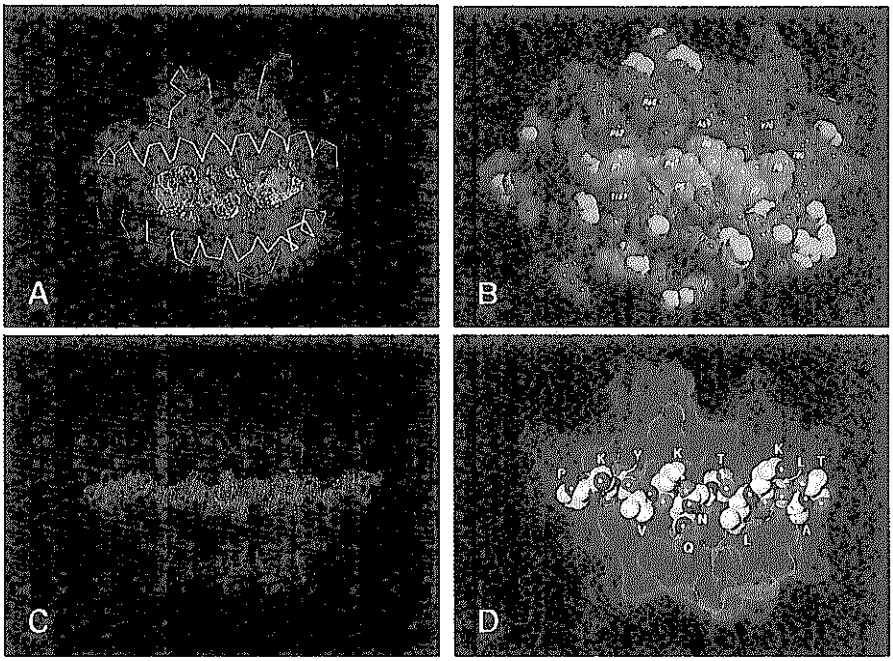

Fig. 2. Peptides in the clefts of class I and class II MHC proteins. Views from the top looking down into the clefts are shown. **A.** van der Waals surface representation of the mixture of peptides in the cleft of HLA-A2 (3). **B.** A single peptide complex of the influenza nucleoprotein peptide 91–99 in the cleft of HLA-A68 (6, 7). Orange atoms are the peptide; red are polymorphic MHC residues. **C.** van der Waals surface representation of the mixture of peptides in the cleft of HLA-DR1 (4). **D.** A single peptide complex of the influenza hemagglutinin peptide 306–318 in the cleft of HLA-DR1 (7).

Ig-like domains (α3 and β2-microglobulin for class I; α2 and β2 for class II) (Fig. 1). The membrane distal non-Ig domains form a superdomain that consists of a platform of 8 β-strands, four from each domain; on top of this platform are two α-helices bowed upwards, one from each of the domains. Peptides are bound for presentation to the immune system in the cleft between the helices and the floor of β-strands. The cleft is lined by polymorphic amino acids. Although the polymorphic residues are spread throughout the approximately 90 amino acids of each of the α1 and α2 (class I) or α1 and β1 (class II) domains, when the molecule folds, all of the polymorphic amino acids end up in the binding cleft.

By good fortune, the first molecule studied, the class I glycoprotein HLA-A2, crystallized with a mixture of peptides in the cleft, and thus the site of foreign peptide binding was identified (Fig. 2A) (3). In the class I protein the peptide

density ended within the structure of the molecule, so that the entire peptide was within the cleft. In cross-section, pockets in the class I molecule were evident. These pockets were filled with electron-dense material, the side chains of individual amino acid residues. The atomic details of the MHC class I molecule's interactions with peptide antigens were first observed in the structure of HLA-B27 (5), the MHC molecule closely associated with ankylosing spondylitis, and later in complexes with single peptides (6, 7). The NH_3^+-terminus points downward and is very deep within a pocket in the structure, where it is fixed by a network of hydrogen bonds derived from tyrosine residues and a water molecule. The side chain of the first amino acid (P1) points upward, but the side chains of P2 and P3 are in pockets (Fig. 2B) (5–7). The carboxyl end, P9, has its side chain rather than its charged groups deep in the pocket, and the carboxylate is relatively superficial. The peptide is largely fixed by hydrogen bonds that cluster at the two charged termini (reviewed in ref. 8). A kink at P3 and P4 allows the center of the peptide to come out of the cleft, and the prominent side chains of P4, P5, P6 and P8 point upward, while P7 is sometimes in a pocket but in other cases P7 points upward. Those side chains that point upward are recognition sites for the effector cells of the immune system, the T cells. The peptide is bound very tightly by additional hydrogen bonds to backbone atoms near the termini, with additional stabilizing force provided by the amino acid side chains in the pockets. These class I MHC/peptide complexes, the transplantation antigens, function in the generation of a cytotoxic T-lymphocyte response.

Class II molecules had already been isolated and a few years later were solubilized and crystallized, and the structure was solved (4). Superposition of the alpha carbon diagrams revealed how similar the class I and class II molecules are. However, a number of important differences in class II allowed the peptide to go through the cleft. The peptides in class II, unlike those in class I, do not terminate within the structure, but extend out both ends (Fig. 2C) (4). They have no kink but exist as an extended twisted peptide structure (a polyproline-like helix). One very prominent density near the N-terminus goes down into a deep pocket.

Other crystal structures involving class II proteins were important in the structure determination, those of HLA-DR1 complexed with a single peptide derived from the influenza virus hemagglutinin and with the superantigen, staphylococcal enterotoxin B (SEB) (9, 10) as well as later with toxic shock syndrome toxin, TSST-1. The DR1/hemagglutinin 306–318 structure (Fig. 2D) (9, 10) showed that the peptitle is fixed in the cleft by an extensive system of hydrogen bonds between

the side chains (or in a few cases backbone atoms) of class II amino acid residues and the backbone of the peptide structure. The deep pocket, particularly important in the fixation of peptides to class II molecules, is occupied by a tyrosine residue in the hemagglutinin peptide (Y308), called P1. Four other pockets in DR 1 fix the side chains of P4, 6, 7 and 9. The residues that flank P1 and P9, i.e. P-1, P-2, P10 and P11, are within the structure so that the cleft itself accommodates 13 residues (Fig. 2D). As in the case of class I, some of the binding energy is provided by the interactions of the peptide chains with the pockets within the MHC molecule but, in the case of class II, the hydrogen bonds to MHC backbone atoms are spread all along the structure, rather than being clustered at the ends (reviewed in ref. 11). These class II MHC/peptide complexes, the Ir gene products, function in the generation of the humoral immune (antibody) response, as well as in a cytotoxic response particularly important in some autoimmune diseases.

Study of the peptides bound to the two classes of molecules clearly revealed the size difference and the fact that, although each peptide epitope bound to class I is limited to a single size, those bound to class II can vary in length at both ends since the cleft which binds them is open (reviewed in ref. 12). Most interestingly, however, the vast majority of peptides bound to both classes of molecules are derived from self proteins. The mechanism of peptide binding in both cases, i.e. the nature of the hydrogen bonds and the "degeneracy" of most of the pockets with respect to the amino acid side chains bound evolved to allow the binding of a universe of foreign (as well as self) peptides to a large but limited number of MHC molelcules.

UTILIZING THE GENETIC AND STRUCTURAL INFORMATION ON CLASS II MHC MOLECULES TO STUDY AUTOIMMUNE DISEASES

This structural information will be of great utility in the study of the genesis and maintenance of autoimmune diseases, a situation in which tolerance to self-peptides bound to class II MHC molecules is abrogated. Most of these autoimmune diseases are genetically linked to particular alleles of the class II molecules, and much sequence information has also accumulated on the polymorphisms in these genes. In the case of multiple sclerosis (MS), linked to the gene DRB1*1501 (that encodes the protein allotype HLA-DR2b), it has long been speculated that viruses and/or other infectious agents could trigger the disease. This hypothesis has been supported by epidemiological studies (for example, of an outbreak of multiple sclerosis in the Faroe Islands after their occupation by British soldiers during

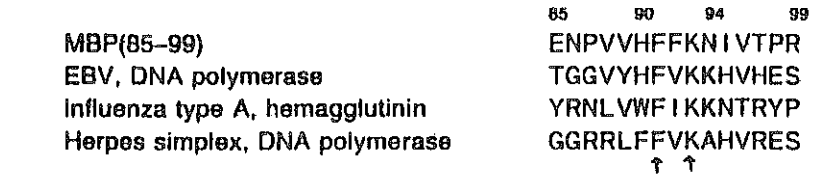

Fig. 3. Viral mimics of myelin basic protein (MBP) peptide 85–99 (13). All of these peptides activated an DR2-restricted MBP-specific T-cell clone from a multiple sclerosis patient. Note that only residues F91 (P3) and K93 (P5) (arrows) are conserved. Amino acids are represented by the single letter code.

World War II). The target of attack is thought to be a normal myelin sheath protein, possibly myelin basic protein (MBP). By using knowledge of how an immunodominant epitope of MBP (MBP85-99) is bound to DRB1*1501 and which of its amino acid side chains is important in contacting the T-cell receptor, seven viral and one bacterial peptide were identified that are effective mimics (13). Two principles were important in the identification. First, the pocket in the MHC molecules that bind amino acid side chains are "degenerate" with respect to the precise side chains that can be accommodated. However the side chains in the peptide that contact the T-cell receptor are nearly completely conserved. Using these principles, effective functional mimics were identified that have no obvious sequence similarity to the MBP peptide, except for the conservation of two amino acid side chains that are essential for contacting the T-cell receptor (Fig. 3). This phenomenon can be termed "degenerate molecular mimicry". These viral and bacterial mimics may provide an important clue to the initiation of MS or to the occurrence of relapses, but it should be emphasized that much more work remains, particularly using materials from patients with very early MS, to prove which, if any, of these viral and bacterial peptides are actually important in the disease. If viral or bacterial mimics can be definitively identified as causative agents, then many therapeutic opportunities are presented.

In a second study of pemphigus vulgaris (PV) (a blistering disease of the skin), linked to DRB1*0402 (HLA-DR4 Dw10), the genetic and structural information on class II MHC molecules was used to predict several peptides from desmoglein 3 (an adhesion molecule important in the adherence of the epidermis to the underlying dermis), one or several of which might be important in the initiation of disease. There are 22 known subtypes of DR4. The subtype linked to pemphigus has a *negatively* charged pocket to accommodate the P4 side chain, which must therefore be positively charged. All of the other subtypes have *positively*

charged P4 pockets accommodating negatively charged side chains. This and other features of peptide binding to DR4 molecules allowed the prediction of seven peptide epitopes of desmoglein 3 (14, 15). Examination of these peptides by proliferation of peripheral blood cells from PV patients has resulted in the identification of a desmoglein 3 peptide that is presented by the disease-associated DR4 molecule. This study offers opportunities similar to those described for MS.

In addition, the techniques and concepts developed in the course of these studies are being used by many others to study peptides bound to MHC alleles linked to other autoimmune diseases and, in addition, peptides and other substance presented by MHC molecules that may be important as vaccines to stimulate anti-tumor responses or responses against important infectious pathogens such as malaria and tuberculosis. It illustrates how the most fundamental investigations are important in illuminating human clinical problems and opening therapeutic possibilities.

ACKNOWLEDGMENT

This work has been generously supported by the National Institutes of Health and by the Howard Hughes Medical Institute. The number of students and postdoctoral fellows as well as collaborators who have been involved in this research over the past 25 years is too large to permit individual enumeration, but we are extremely grateful to each of them. Their names appear in the reviews cited above.

REFERENCES

1. Ploegh HL, Orr HT, Strominger JL. Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell* 1981; 24: 287–99.
2. Kaufman JF, Auffray C, Korman AJ, Shackelford DA, Strominger JL. The class II molecules of the human and murine major histocompatibility complex. *Cell* 1984; 36: 1–13.
3. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2 and 2. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens *Nature* 1987: 329: 506–12, 512–8.
4. Brown JH, Jardetsky TS, Gorga JC et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 1993: 364: 33–9.
5. Madden DR, Gorga JC, Strominger JL, Wiley DC. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature* 1995: 353: 321–5.
6. Silver ML, Guo HC, Strominger JL, Wiley DC. Atomic structure of a human MHC molecule presenting an influenza virus peptide. *Nature* 1992: 360: 367–9.
7. Madden DR, Garboczi DN, Wiley DC. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* 1993: 75: 693–708.
8. Madden DR, Wiley DC. Peptide binding to the major histocompatibility complex molecules. *Curr Biol* 1992: 2: 300–4.
9. Stern LJ, Brown JH, Jardetzky TS et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994: 368: 215–21.
10. Jardetzky TS, Brown JH, Gorga et al. Three-dimensional structure of human class II histocompatibility molecule complexed with surperantigen. *Nature* 1994: 368: 711–8.

11. Stern LJ, Wiley DC. Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure* 1995: 2: 245–51.
12. Urban RG, Chicz RM, Vignali DAA, Strominger JL. The dichotomy of peptide presentation by class I and class II MHC proteins. *Chem Immunol* 1993: 57: 197–234.
13. Wucherpfennig KW, Strominger JL. Molecular mimicry in T cell mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell* 1995: 80: 695–705.
14. Wucherpfennig KW, Strominger JL. Selective binding of self-peptides to disease associated MHC molecules: a mechanism for MHC linked susceptibility to human autoimmune diseases. *J Exp Med* 1995: 181: 1597–601.
15. Wucherpfennig KW, Yu B, Bhol K, Ahmed AR Strominger JL. Structural basis for MHC linked susceptibility to autoimmunity: charged residues of a single MHC binding pocket confer selective presentation of self-peptides in pemphigus vulgaris. *Proc Natl Acad Sci U S A* 1995: 92: 11935–9.