Atomic structure of a human MHC molecule presenting an influenza virus peptide

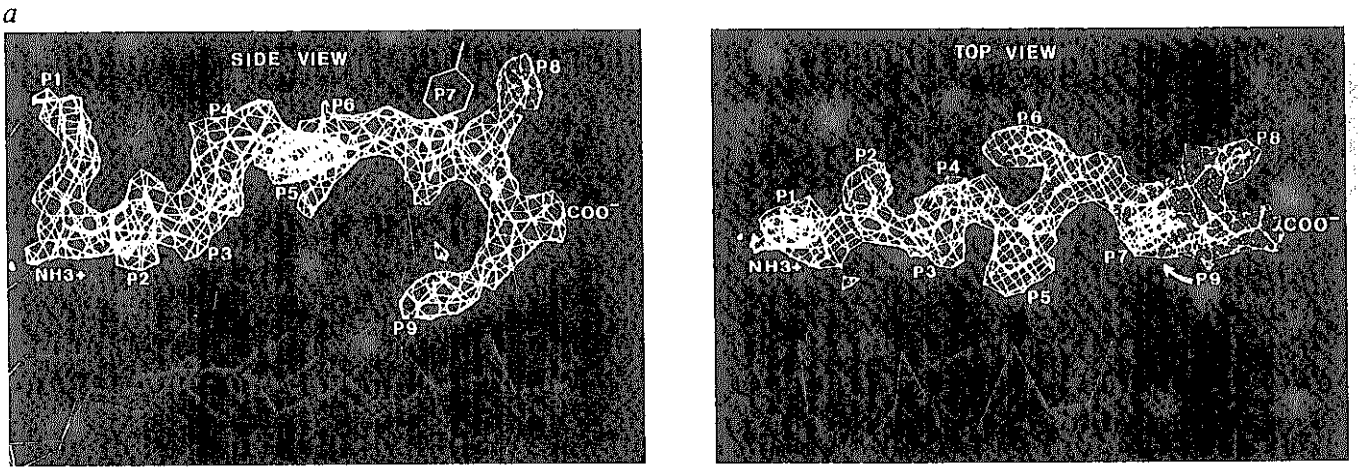
Michael L. Silver*, Hwai-Chen Guo*, Jack L. Strominger* & Don C. Wiley†

* Department of Biochemistry and Molecular Biology and
† Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

INFECTION by influenza virus results in the stimulation of cytotoxic T lymphocytes specific for killing virally infected cells[1]. Specificity is provided by clonally distributed, hypervariable T-cell receptors on cytotoxic T lymphocytes which react with peptide fragments that are derived from viral proteins expressed in the cytoplasm and 'presented' on the surface of infected cells, bound to class I histocompatibility glycoproteins[2]. Here we describe the structure of the complex between the human class I histocompatibility glycoprotein HLA-Aw68 and the influenza virus nucleoprotein peptide Np 91–99 as determined by X-ray cryocrystallography. Residues at both ends of the peptide are substantially buried in the peptide binding-site, whereas those in the middle of the peptide, P4 to P8, are predominantly exposed and could be recognized directly by T-cell receptors. The extended conformation of the bound viral peptide is remarkably similar to that of a collection of endogenous peptides with a different sequence motif bound to another human allele, HLA-B27[3–5]. The structure defines in atomic detail the antigenic surface constructed of major histocompatibility complex and viral peptide atoms that is recognized by T-cell receptors.

Crystals of HLA-peptide complex were grown from HLA-Aw68, isolated from human lymphoblastoid cells[6,7] and reconstituted *in vitro*[8] with the synthetic peptide Np 91–99 (Lys-Thr-Gly-Gly-Pro-Ile-Tyr-Lys-Arg)[9]. The structure was determined to 2.8 Å resolution from a single crystal frozen at −160 °C and refined to an *R*-factor of 18.5% with good geometry (see Fig. 1 legend). Electron density corresponding to Np 91–99 (Fig. 1*a*) was clearly interpretable with the exception of the side chain of tyrosine (P7) which appears to extend into solution and to be disordered.

The peptide conformation and mode of binding to the HLA groove are essentially the same for Np 91–99 binding to HLA-Aw68 as for multiple endogenous self-peptides bound to HLA-B27[3–5]. When the HLA-Aw68 and HLA-B27 molecules are superimposed, the atomic models of their bound peptides are nearly coincident (Fig. 1*b*). In both cases the peptide is extended with a kink at P3 and P4. Peptide side chains P2, P3 and P9 face into the major histocompatibility complex (MHC) groove, P1, P4, P8 face away towards the T-cell receptor (TCR), and P5 and P6, although accessible to a TCR, point back and forth across the top of the groove (Fig. 1*a*). The side chain at P7 points toward the TCR in the Aw68-Np 91–99 complex, but toward the groove in HLA-B27. Thus, the overall mode of



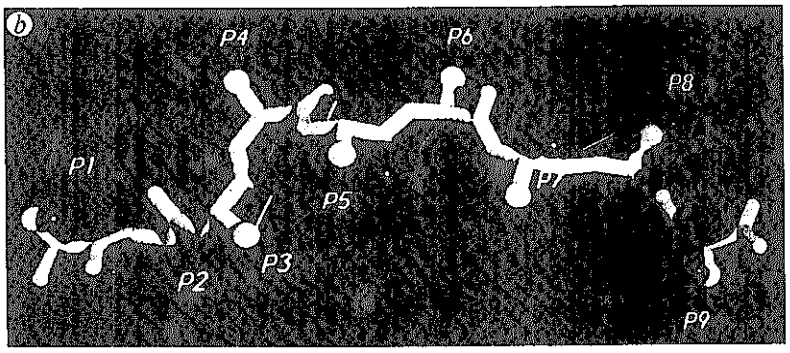


FIG. 1 The structure of influenza virus peptide Np 91–99 as bound to HLA-Aw68. *a*, Top and side views of the 2.8 Å difference electron density with peptide model. The α_1 domain α-helix runs from left to right. *b*, Comparison of the peptide conformations of Np 91–99 (red) bound to HLA-Aw68 and the model from a collection of endogenous peptides bound to HLA-B27 (white)[3–5]. The main-chain coordinates of HLA-Aw68 and HLA-B27 both without peptide coordinates were superimposed by FITATOM, a locally modified version of the program EXIFIT written by A. D. MacLachlan. The transformation matrix obtained was then used to superimpose the peptide models.

METHODS. The crystal structure of HLA-Aw68 complexed with Np 91–99 was determined by molecular replacement using 2.1 Å refined coordinates from a crystal of HLA-Aw68 complexed with a collection of endogenous peptides[10]. X-ray data to 2.8 Å were collected from a single crystal, frozen in a film of buffer in a wire loop[14] at −160 °C, using a Xentronics detector[15] and processed with BUDDHA software[16]. Subsequent calculations used AGROVATA and ROTAVATA[17] (R_{merge} = 10% to 2.8 Å) and XPLOR[18]. Initial molecular replacement solution plus positional minimization of HLA atoms revealed clear density for peptide main chain and four of the seven side chains (Thr P2, Pro P5, Ile P6, Arg P9) in an $F_o - F_c$ map. A peptide model was built into the clear density. Subsequent cycles of refinement brought up electron density interpretable for two of the remaining three side chains (Lys P1, Lys P8) (r.m.s. deviation in bonds = 0.016 Å, r.m.s. deviation in angles = 3.56°). Np 89–101 had been identified as an influenza epitope restricted to HLA-Aw68[9]. Np 91–99 was chosen from a series of truncated analogues because it produced the maximum in yield in *in vitro* complex formation[19]. *a*, Generated with FRODO[20] and *b*, with RASTER3D, a program written by D. Bacon and W. F. Anderson.

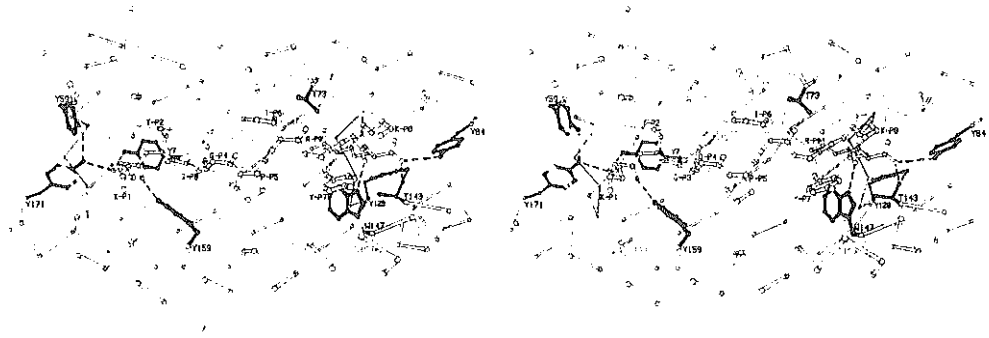


FIG. 2 Peptide binding interactions. Peptide in orange. HLA residue side chains containing contacts within 4 Å of peptide are shown. Conserved human class I MHC residues in blue (143 is Ser and 147 is Leu in one sequence). Variable MHC residues and water molecules in red. Dashed lines show potential hydrogen bonds from Np 91–99 to HLA-Aw68, blue dashes to conserved class I MHC residues, red dashes to variable class I MHC residues either directly, or bridged by water. Selection of conserved and polymorphic residues based on ref. 21. MHC residues of low variability are 63, 99, 150, 152 and 167. Stereo figure generated with ORTEP[22].

peptide binding is extremely similar, although HLA-Aw68 and HLA-B27 differ by 12 amino acids in the peptide-binding site and the sequence motif of peptides that bind to the two alleles also differs[4,10].

Atomic contacts at the two ends of the peptide predominate in binding Np 91–99 to HLA-Aw68. Residues P1, P2, P3 and P9 are substantially buried in the antigen-binding cleft. In contrast, the central peptide residues P4 to P7 make no hydrogen bonds to the HLA molecule and P4 and P5 do not even make van der Waal contacts (Fig. 2). Eleven MHC residues contact the first two peptide amino acids P1 and P2 and eight other MHC residues contact the last peptide amino acid P9, whereas all the rest of the peptide, P3 to P8, is in contact with only 10 MHC residues. In the centre of the peptide where direct contact to HLA is sparse, water molecules provide a hydrogen-bonding bridge to HLA (Fig. 2).

Conserved side chains of the HLA molecule hydrogen bond to main-chain polar atoms at both ends of the peptide: HLA residues Tyr 7 and Tyr 171 to the peptide N terminus, Tyr 159 to the P1 carbonyl oxygen, Tyr 84 and Thr 143 to the C-terminal carboxylate, and Trp 147 to the P8 carbonyl oxygen (Fig. 2). Similar hydrogen-bond networks, constituting conserved peptide terminal binding sites, have been visualized in the complexes of multiple endogenous peptides bound to HLA-B27[3-5] and to HLA-Aw68 (Fig. 2 in ref. 10).

Peptide side chains Thr P2, and Arg P9 are entirely buried in binding cleft pockets composed of polymorphic residues. The side chain of threonine P2 is bound by atoms from five MHC residues (Tyr 7, Met 45, Asn 63, Asn 66, Tyr 99) in the shallow '45' pocket of HLA-Aw68. The side chain of arginine P9 reaches deep into the site, forming salt-bridged hydrogen bonds with the polymorphic residues Asp 77, and Asp 116 (Fig. 2) and van der Waal contacts with a total of six MHC residues (Asp 74, Asp 77, Met 95, Asp 116, Thr 143, Trp 147), four of which (74, 77, 95, 116) are polymorphic in the population. The contacts to Arg P9 are similar to those that Arg P9 would make in HLA-B27 which shares many of the same polymorphic residues as Aw68 in the 'P9 end' of the binding site[3]. The binding of arginine P9 to the negatively charged pocket near Asp 74 was predicted from considering the HLA-Aw68 structure and making substituted peptides, as was the accessibility of Lys P8 to cytotoxic T lymphocyte (CTL) recognition[9,11]. If P3 (Gly) had a side chain it would point into a shallow non-polar pocket. In the accompanying paper[10] HLA-Aw68 is shown to prefer peptides with Val or Thr at P2 and Lys or Arg at P9, as well as a non-polar side chain at P3. That is consistent with the observations here where the two 'anchors' P2 and P9 of Np 91–99 are completely sequestered by MHC atoms.

Six of the nine peptide residues, P1, P4, P5, P6, P7 and P8 remain solvent accessible in the complex with HLA-Aw68 and could be recognized directly by the T cell (Fig. 3). Most of the accessible atoms are from the central residues of the peptide (red in Fig. 3). Six of the peptide's polar groups are accessible with the potential to form hydrogen bonds directly to a TCR, three from side chains Lys P1, Tyr P7 and Lys P8 and three from the main chain (yellow bonds in Fig. 3).

The structure of the complex between the influenza virus peptide Np 91–99 and the human class I histocompatibility glycoprotein HLA-Aw68 shows a class I molecule holding a peptide predominantly by its ends, stretched out so that most of its sequence can be 'read' by TCR (Fig. 4). This makes possible the detection of foreign gene products within a host cell by T-cell surveillance of class I MHC/peptide complexes. Six of the nine peptide residues could be contacted directly by TCR, a number sufficient to discriminate between a very large number of sequences. (A mechanism for presenting an even larger number of sequences, by insertion of additional peptide residues, is reported in the accompanying paper[10].) No large conformational changes were found between HLA-Aw68 complexed with a single viral peptide, Np 91–99, and HLA-Aw68 complexed with multiple endogenous peptides[10], suggesting in this case that peptide induced conformational changes of the HLA molecule

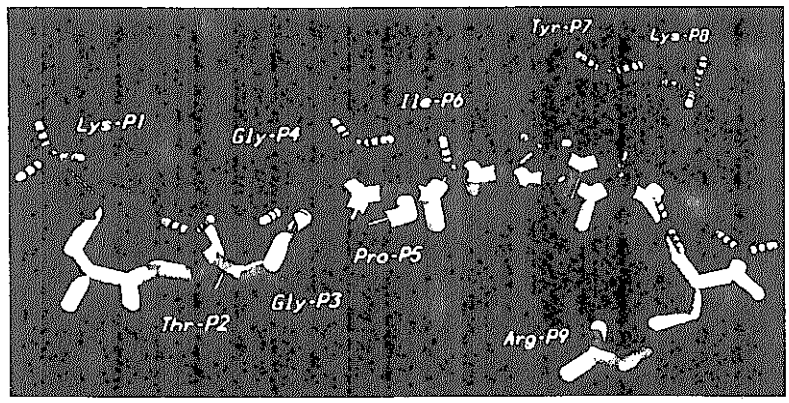


FIG. 3 Influenza virus nucleoprotein peptide Np 91–99 showing atoms accessible to solvent only (dark red) accessible to contact by TCR (orange), or buried by contact with HLA-Aw68 (white). Parts of residue P1 and P4–P8 are accessible to TCR. The solvent accessibility was determined using the program ACCESS written by M. D. Handschumacher and F. M. Richards with a 1.4 Å probe. TCR accessibility was approximated with a 1.9 Å probe representing an average van der Waals radius of TCR atoms. Yellow striped bands represented potential hydrogen bonds directly to TCR (on orange enlarged atoms) or to TCR through water molecules only (on red atoms). Figure generated with RASTER3D. (A total of 1,098 Å^2 of the solvent-accessible surface of the free Np 91–99 peptide is buried in the complex with HLA-Aw68 and 711 Å^2 of the HLA-Aw68 solvent-accessible surface is buried.)

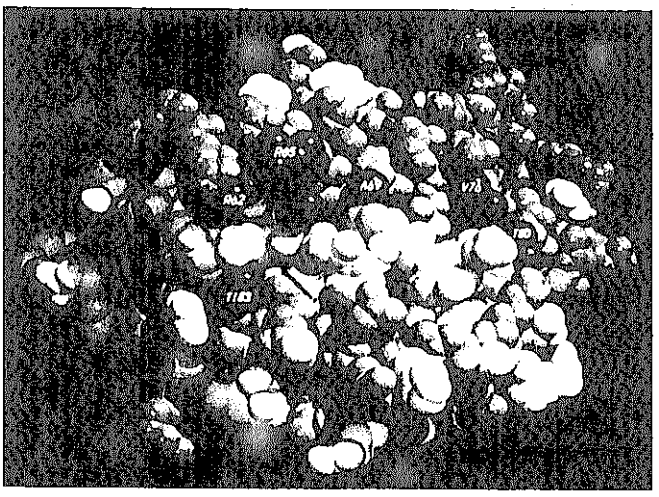


FIG. 4 The 'antigenic' surface composed of Np 91–99 peptide (orange) and MHC atoms (blue, conserved; red, polymorphic; light blue, not conserved or polymorphic). N-terminal of peptide is to the left, α_1 domain α-helix top, α_2 domain α-helix bottom. Figure generated with RASTER3D. (Of the 12 polymorphic residues facing into the binding site, 8 contact the peptide directly (9, 45, 66, 70, 74, 77, 95, 116) and four do not (67, 97, 114, 156), but 11 of the 12 (66 excluded) are nevertheless completely buried by the bound peptide. These polymorphic positions must therefore, as anticipated[23], have their primary effect on T-cell recognition of HLA-Aw68 through the choice of peptides that can bind. Of the six polymorphic residues that face more directly toward solvent (62, 65, 69, 76, 80, 163), four also contact the peptide (62, 69, 80, 163) but all have atoms accessible to direct recognition by the TCR and therefore represent polymorphism recognizable by TCRs in the presence or absence of peptide.)

may not be a major factor in the creation of novel antigenic surfaces recognized by T cells. On the basis of the number of atomic contacts, Np 91–99 appears to be bound to HLA-Aw68 predominantly by two main features of the MHC molecule: (1) conserved MHC residues hydrogen bond to the peptide termini; (2) polymorphic MHC residues bury the two 'anchor' peptide side chains. Although both of these sets of interactions would also provide for the peptide-dependent stabilization of the MHC molecule, only the peptide termini binding sites are conserved in class I histocompatibility antigen sequences. The overall mode of peptide binding observed here seems to be a general mechanism for class I MHC presentation now visualized in three human alleles and one murine allele: HLA-B27[3–5], HLA-Aw68[10], HLA-A2[12] and H-2K^b(refs 13, 24). □

Received 4 August; accepted 2 October 1992.

1. Townsend, A. & Bodmer, H. *A. Rev. Immun.* **7**, 601–624 (1989).
2. Brodsky, F. M. & Guagliardi, L. E. *A. Rev. Immun.* **9**, 707–744 (1991).
3. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **253**, 321–325 (1991).
4. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **253**, 326–329 (1991).
5. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Cell* **70**, 1035–1048 (1992).
6. Turner, M. J. *et al. J. biol. Chem.* **250**, 4512–4519 (1975).
7. Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *J. molec. Biol.* **188**, 205–210 (1986).
8. Silver, M. L., Parker, K. C. & Wiley, D. C. *Nature* **350**, 619–622 (1991).
9. Cerundolo, V., Tse, A. G. D., Salter, R. D., Parham, P. & Townsend, A. *Proc. R. Soc.* **244**, 169–177 (1991).
10. Guo, H-C., Jardetzky, T. S., Lane, W. S., Strominger, J. L. & Wiley, D. C. *Nature* **380**, 364–366 (1992).
11. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **342**, 692–696 (1989).
12. Saper, M. A., Bjorkman, P. J. & Wiley, D. C. *J. molec. Biol.* **219**, 277–319 (1991).
13. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. *Science* **257**, 919–927 (1992).
14. Teng, T-Y. *J. appl. Crystallogr.* **23**, 387–391 (1990).
15. Durbin, R. M. *et al. Science* **232**, 1127–1132 (1986).
16. Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Crystallogr.* **20**, 235–242 (1987).
17. Fox, G. C. & Holmes, K. C. *Acta Crystallogr.* A**34**, 517 (1966).
18. Brünger, A. T. XPLOR (version 2.1) Yale University, New Haven, (1990).
19. Silver, M. L. thesis, Harvard University (1992).
20. Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
21. Parham, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 4005–4009 (1988).
22. Johnson, C. K. ORTEP Manual (Oak Ridge National Laboratory, Tennessee, 1965).
23. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
24. Zhang, W., Young, A. C. M., Imarai, M., Nathenson, S. G. & Sacchettini, J. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8403–8407 (1992).

ACKNOWLEDGEMENTS. We acknowledge the contributions made by T. P. J. Garrett, M. A. Saper, and P. J. Bjorkman to the determination of the structure of HLA-Aw68 to 2.6 Å resolution[11]. We thank J. Gorga for supervising the protein purifications and A. Haykov, K. Svenson and R. Crouse for technical assistance. H.-C.G. is a Cancer Research Institute fellow. J.L.S. and D.C.W. acknowledge support from the NIH. D.C.W. is supported by the Howard Hughes Medical Institute.