# Specificity pockets for the side chains of peptide antigens in HLA-Aw68

**T. P. J. Garrett*, M. A. Saper*, P. J. Bjorkman†, J. L. Strominger & D. C. Wiley*‡**

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

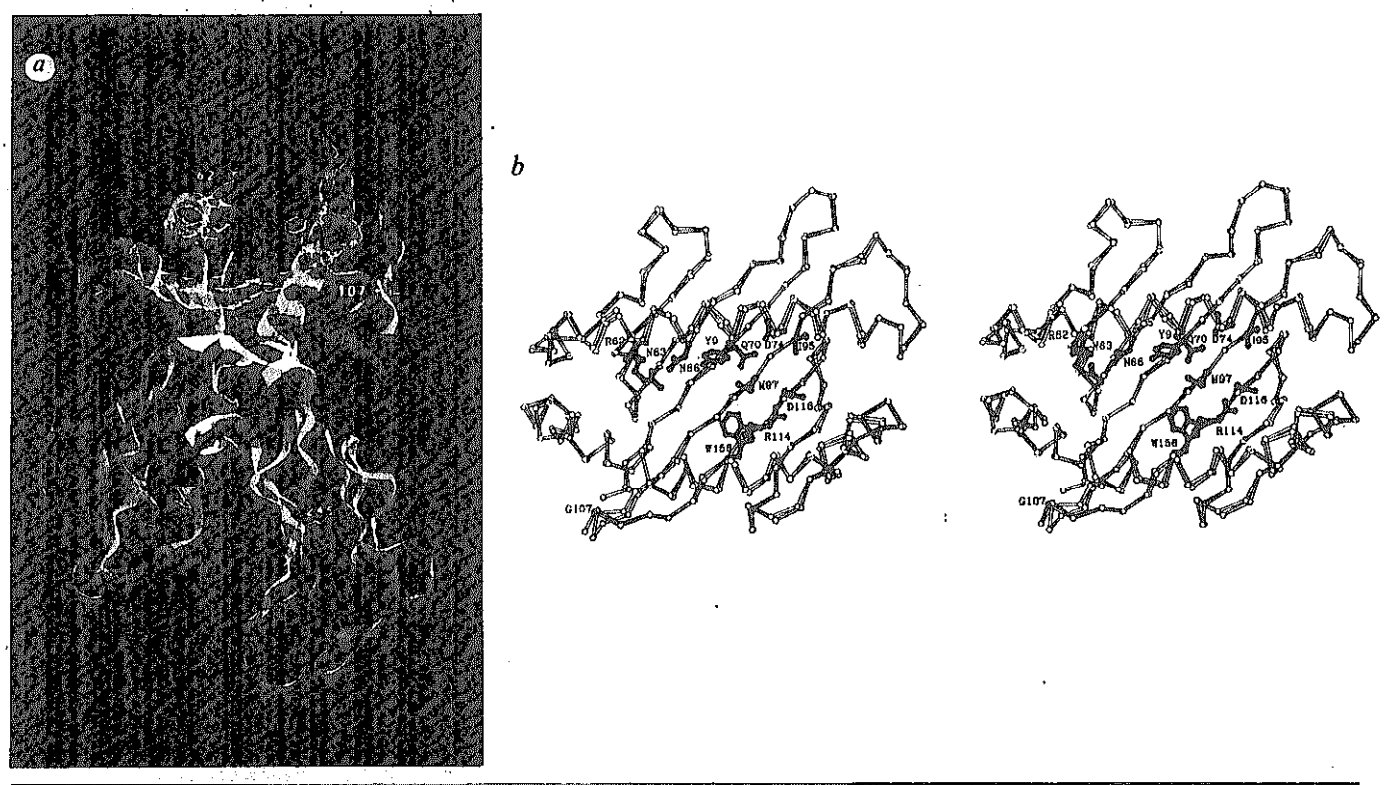
**WE have determined the structure of a second human histocompatibility glycoprotein, HLA-Aw68, by X-ray crystallography and refined it to a resolution of 2.6 Å. Overall, the structure is extremely similar to that of HLA-A2 (refs 1, 2; and M.A.S. *et al.*, manuscript in preparation), although the 11 amino-acid substitutions at polymorphic residues[3,4] in the antigen-binding cleft[2] alter the detailed shape and electrostatic charge of that site. A prominent negatively charged pocket within the cleft extends underneath the α-helix of the $\alpha_1$-domain, providing a potential subsite for recognizing a positively charged side chain or peptide N terminus. Uninterpreted electron density, presumably representing an unknown 'antigen(s)', which seems to be different from that seen in the HLA-A2 structure[1], occupies the cleft and extends into the negatively charged pocket in HLA-Aw68. The structures of HLA-Aw68 and HLA-A2 demonstrate how polymorphism creates and alters subsites (pockets) positioned to bind peptide side chains, thereby suggesting the structural basis for allelic specificity in foreign antigen binding.**

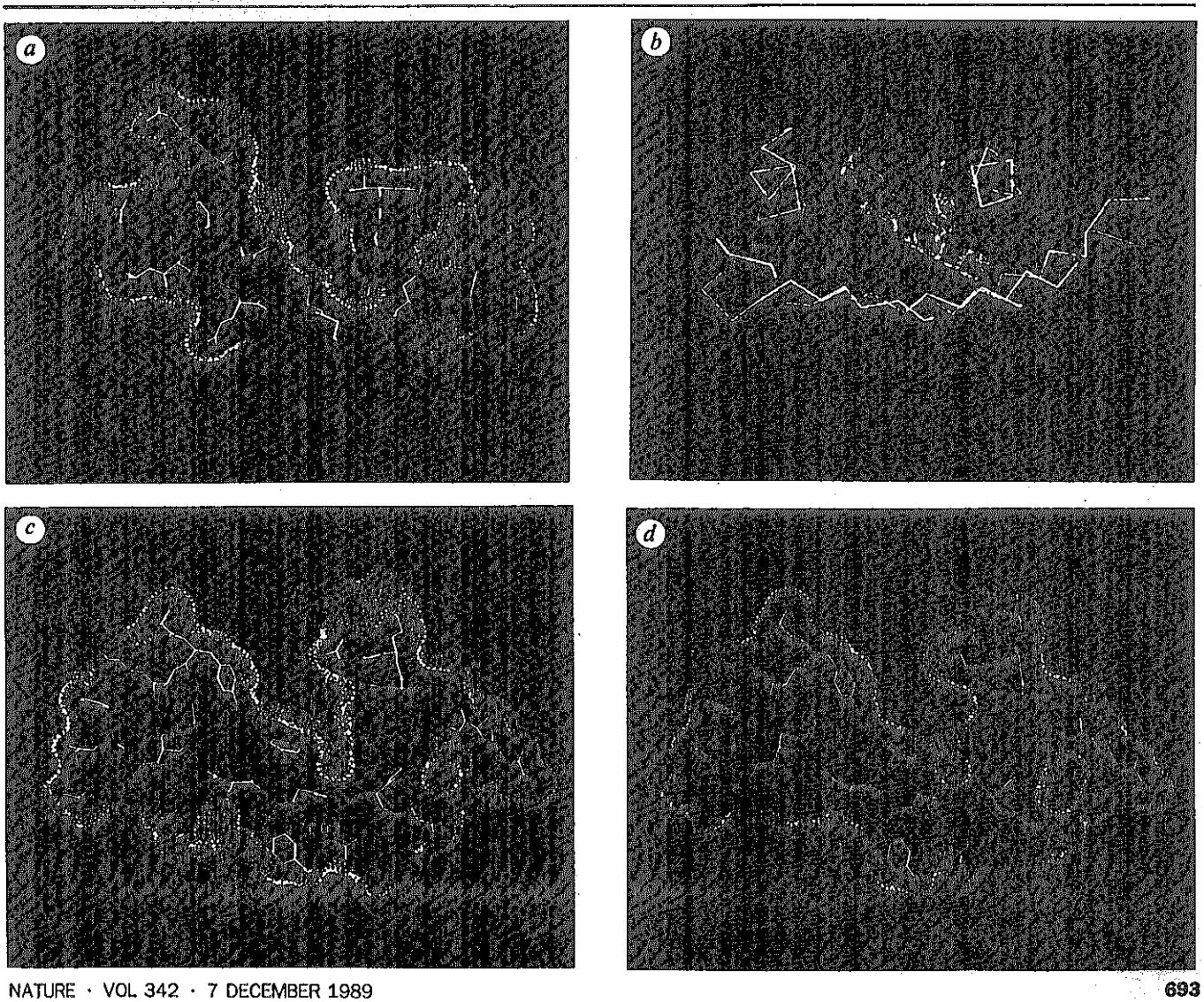
HLA-Aw68 (formerly HLA-A28) is the product of one allele of the polymorphic HLA-A locus of the major histocompatibility complex (MHC). Like other class I histocompatibility glycoproteins HLA-Aw68 seems to form complexes with peptides derived from internal degradation of self and nonself antigens[5–8]. These complexes are recognized on cell surfaces by cytotoxic T lymphocytes (CTL) leading to the elimination of virally infected or histoincompatible cells. The extensive polymorphism of histocompatibility glycoproteins is concentrated at the putative peptide-binding cleft[2,9], indicating that the variability was selected to generate diversity in peptide binding and T-cell receptor recognition[2]. Individual class I and class II histocompatibility glycoproteins interact with an extremely broad range of peptides of diverse sequences[10–14], yet exhibit selectivity by failing to interact with many peptides. This failure to interact probably

† Present address: Division of Biology, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California, USA.
‡ To whom correspondence should be addressed.

FIG. 1 Comparison of HLA-Aw68 and HLA-A2. *a*, Main-chain ribbon diagram of HLA-Aw68 with the 13 side chains of residues that differ from those of HLA-A2. Ten differences between HLA-A2 and HLA-Aw68, respectively, (F9Y, E63N, K63N, H70Q, H74D, V95I, R97M, H114R, Y116D and L156W) are in the binding cleft. One difference, G62R, faces up from the $\alpha_1$-domain α-helix (labelled 62). W107G (labelled 107) is on a loop in the $\alpha_2$-domain and A254V (labelled 245) is on the $\alpha_3$-domain. $\beta_2$-microglobulin is red. *b*, Stereoview of HLA-Aw68 and HLA-A2 showing the α-carbon atoms of the $\alpha_1$- and $\alpha_2$-domains superimposed. Side-chain differences in HLA-Aw68 are in solid lines. Root-mean-square differences between atomic positions of α-carbons (with $B \leq 35$ Å$^2$) in HLA-Aw68 and HLA-A2: domains $\alpha_1$ and $\alpha_2$, 0.45 Å; $\alpha_3$-domains, 0.27 Å; and $\beta_2$-microglobulin, 0.30 Å; Standard deviation (s.d.) of overall coordinate errors in each structure are estimated at 0.3 Å (for further details, see refs 18, 19, and M.A.S. *et al.*, manuscript in preparation). (In the site only a few small shifts in main-chain positions (over 2 s.d. of the error) are observed, one in the α-helix of $\alpha_1$-domain at residues 61–71, another in the α-helix of $\alpha_2$-domain in the region 148–156, and a third in a β-strand at positions 97, 98 and 101 (Fig. 1*b*).) For details of structure determination, see ref. 18. Briefly, the initial model was based on the structure of HLA-A2 at 3.0 Å. Side-chains that differed between the two molecules were fit in electron density maps calculated with coefficients $F_O$ (Aw68) and $F_O$ (Aw68) − $F_O$(A2) and phases calculated from the HLA-A2 model. The structure was refined using the program, XPLOR[30,31]; current *R* factor = 17.9% for all data; $I > 1\sigma$; 6–2.6 Å; 7 water molecules; r.m.s. deviations from ideal bond lengths, 0.016 Å; angles, 3.5°. When residues that differ between HLA-Aw68 and HLA-A2 were omitted, difference maps showed clear electron density for the ten residues pointing into the site and for 245. Density for Gly 107 and the side-chain of Arg 62 appeared weak owing to disorder. An analysis of the effect of crystal contacts on the structures of HLA is found in ref. 17. *a*, Prepared with HYDRA (R. Hubbard, unpublished); *b*, Prepared with PLUTO (S. Motherwell, E. Dodson and P. Evans, unpublished). Single-letter amino-acid code.

FIG. 4 Sections perpendicular to the length of the cleft showing different and similar pockets in HLA-Aw68 and HLA-A2 binding sites. *a*, *c*, *d*, Sections through solvent contact surfaces coloured by charge: red, acidic; blue, basic; green, His; yellow, neutral. The cleft appears as a central depression; the $\alpha_1$-domain α-helix is towards the top right, the $\alpha_2$-domain α-helix is at the top left; β-sheet residues form the bottom of the groove. *a*, A pocket at Asp 74 under the $\alpha_1$-domain α-helix of HLA-Aw68. Note red 'dots' on the top of the pocket indicate the surface is formed by the carboxylate oxygens of Asp, 74. *b*, Same view of HLA-Aw68 as in *a* showing extra electron density (red) which fills the cleft, reaching into the '74 pocket'. Picture is sectioned through a blue van der Waal's surface of HLA-Aw68 and has a yellow α-carbon backbone. *c*, The pocket at Met 45 of HLA-Aw68 extends under the $\alpha_1$-domain α-helix. *d*, The pocket at Met 45 HLA-A2 is similar to that in *c*. Note also the essentially identical structure of HLA-Aw68 and HLA-A2 in regions where atoms are labelled in red. (Label: Sequence number, atom type). The extra electron density bound to HLA-Aw68 is not currently interpretable as a single conformation such as an α-helix. Figures constructed using HYDRA (R. Hubbard, unpublished) and MS (ref. 29).


a
b
R82
N63
Y0
Q70
D74
I93
N66
M07
D116
R114
W158
G107


a
b
c
d

provides the basis for allele-specific nonresponsiveness to certain immunological challenges[10,11,15,16]. Here we report the structure of HLA-Aw68 refined to a resolution of 2.6 Å, which when compared with the 2.7 Å resolution structure of HLA-A2 (M.A.S. *et al.*, manuscript in preparation, and ref. 17) suggests mechanisms for both the broad specificity and the allelic selectivity of peptide binding.

The structure of papain-solubilized HLA-Aw68 was determined and refined to 2.6-Å resolution (see ref. 18 for details) from orthorhombic crystals isomorphous to HLA-A2 crystals[19]. The overall structure, excluding the 13 amino-acid differences, is very similar to that of HLA-A2 refined to high resolution from both monoclinic crystals (2.7 Å; M.A.S. *et al.*, manuscript in preparation) and orthorhombic (2.5 Å) crystals[17]. Figure 1*b* shows the $\alpha$-carbon backbone of HLA-Aw68 overlaid on the coordinates of the 2.7 Å resolution monoclinic crystal structure of HLA-A2 and lists root mean square (r.m.s.) deviations in coordinates between domains of the two structures. The average deviations are within the expected experimental error in the coordinates of these structures (Fig. 1*b*, legend).

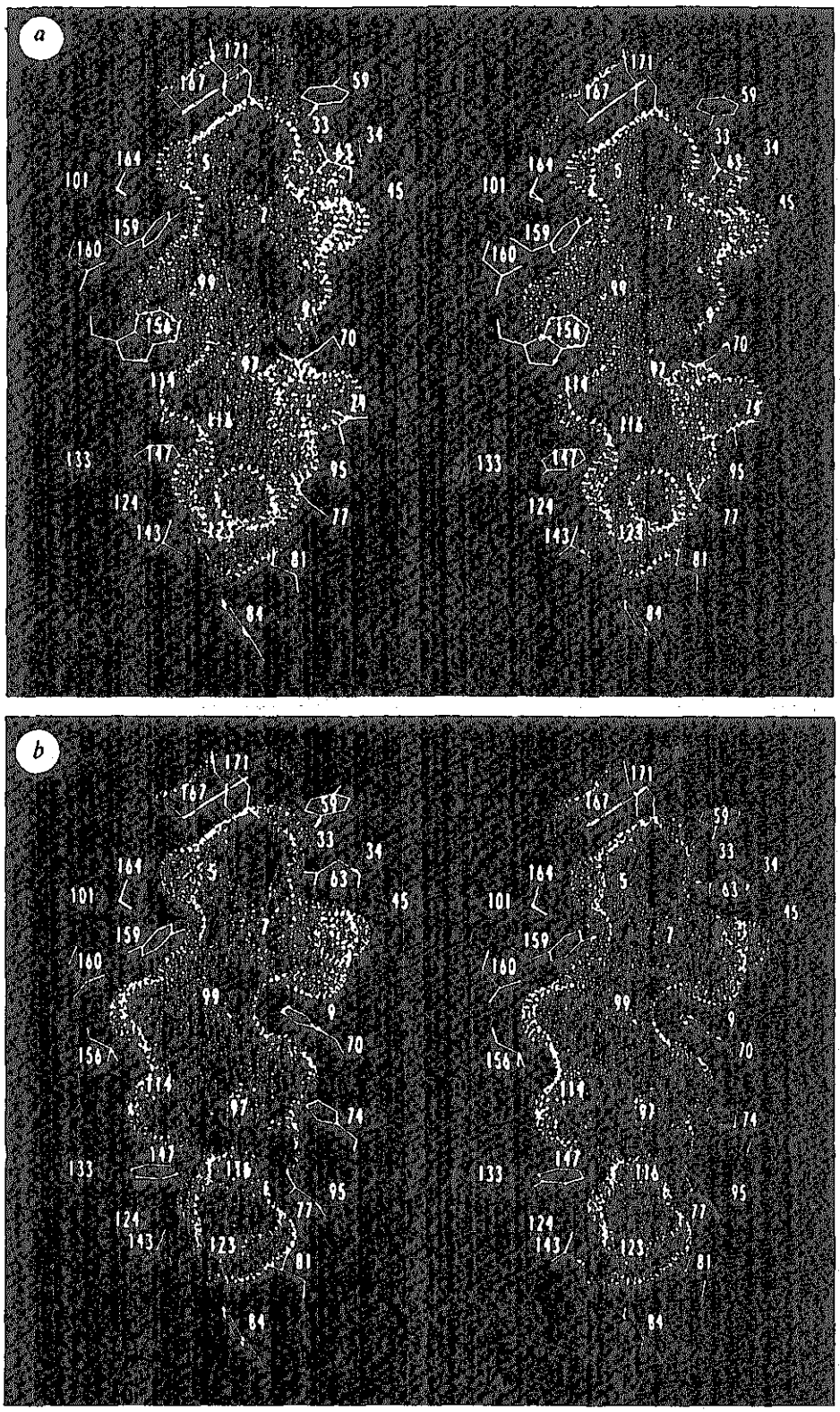



FIG. 3 Top view, comparison of binding site of HLA-Aw68 with HLA-A2. Stereograms show solvent contact surfaces of the clefts coloured by side-chain atom type: red, oxygen; blue, nitrogen; yellow, carbon; purple, sulphur. Side chain colours indicate secondary structure: red side chains, $\beta$-sheet; blue side chains, $\alpha$-helix. *a*, HLA-Aw68 shows prominent new pockets near positions 74 and 116 and a ridge at positions 156 and 70, *b*, HLA-A2 shows Arg 97, Tyr 116, and His 74 filling the pockets. Both clefts have a pocket at 45. *c*, *d* Schematic outline of HLA-Aw68 (*c*) and HLA-A2 (*d*) cleft with pockets labelled. *a*, *b*, Constructed with HYDRA (R. Hubbard, unpublished) and surfaces were calculated using the program MS (ref. 29). Single-letter amino-acid code.

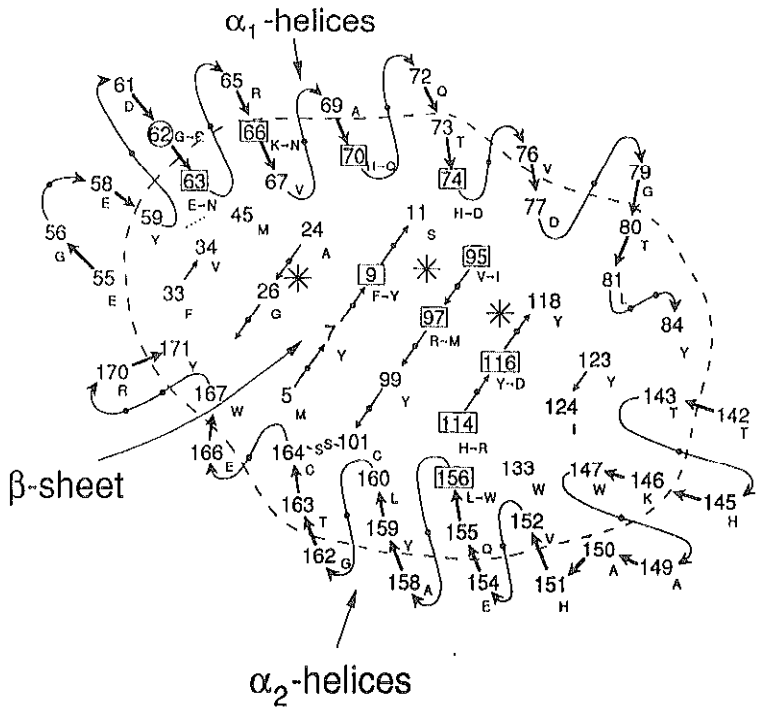



FIG. 2 Diagram of class I binding cleft. Thin arrows (at centre) are $\beta$-sheet residues on the cleft floor. Helical residues are shown at both edges, bold arrows are the top helical surface. Sequence labels are HLA-A2, with HLA-Aw68 shown after a short arrow (62G→R). Residues inside dashed line are ≤4.5 Å from extra electron density and some may contact peptides. Differences between HLA-A2 and HLA-Aw68 are boxed. Stippled residues are conserved in all human class I sequences[9,28]. Asterisks mark pockets described in text. Single-letter amino-acid code.

The principal differences between the structures of HLA-Aw68 and HLA-A2 are in the size and precise location of the side chains of the 13 amino-acid differences (Fig. 1). Ten of these face into the putative peptide-binding cleft: residue 62 points 'up' from the $\alpha_1$-domain $\alpha$-helix in a position where it could be recognized directly by a T-cell receptor[2]; residue 107 is on a loop between $\beta$-strands of the $\alpha_2$-domain; and residue 245 is on the $\alpha_3$-domain at the putative contact site with the accessory molecule CD8 (refs 20, 21). The number of amino acids that differ within the site (10) is approximately equal to the average number of site differences between pairs of sequences within HLA loci (9.7). Therefore the HLA-Aw68–HLA-A2 comparison should provide a representative view of the effect of polymorphic changes within the antigen-binding site. Figure 2 summarizes diagrammatically the location of the 10 differences in the antigen-binding cleft and how they affect the lining of the cleft. For example, the asterisks indicate pockets formed in the checkerboard pattern of $\beta$-sheet side-chains between residues 9, 11, 97, 95 and 74 and between residues 95, 97, 116 and 118 that are opened up in HLA-Aw68 by the substitutions at residues 74, 95, 97, 114, and 116.

Although each individual amino-acid difference in the binding cleft seems to cause only very local structural changes (limited to a few tenths of angstroms in nearest neighbour side chains), the 10 substitutions taken together substantially transform the shape and charge of parts of the cleft but leave other parts unchanged (Fig. 3). Figure 4*a* illustrates the most dramatic difference between the two structures—a deep negatively charged pocket formed below the $\alpha_1$-domain $\alpha$-helix of HLA-Aw68 by the substitution of Asp for His at position 74. Access to this pocket is facilitated by the substitutions of Met for Arg at position 97 and Asp for Tyr at position 116 on the $\alpha_2$-domain $\beta$-sheet (compare Fig. 3*a* and *b*). Figure 4*b* shows that the 'extra electron density' which fills the antigen-binding cleft extends into this pocket. The top surface of the pocket is formed by the carboxylate oxygen atoms of Asp 74 (red dots, Fig. 4*a*), so that this pocket seems suited to bind polar atoms, especially a positively charged side-chain or N terminus of an antigenic peptide. The sequences of other class I glycoproteins indicate that a pocket with this specificity is probably present in several HLA-A and -B molecules. Other prominent alterations in HLA-Aw68 include a deep pocket in the $\alpha_2$-domain $\beta$-sheet caused by the substitution of the small amino acid, aspartic acid, for a large aromatic tyrosine at position 116 (see Fig. 3*a*; asterisks in Fig. 2), and the presence of a ridge across the centre of the bottom of the cleft owing to the bulky substitutions of tryptophan at position 156 and glutamine at position 70 (Fig. 3*a*).

One prominent pocket below the $\alpha_1$-domain $\alpha$-helix remains essentially conserved between the two structures (Fig. 4*c*, *d*). This pocket is a hydrophobic cavity with Met 45 at its base (surrounded by Val 67, Val 34, Gly 26 and Ala 24; (asterisks in Fig. 2, and Fig. 3)). It seems large enough to bind a leucine side chain (M.A.S. *et al.*, in preparation). Comparison of HLA sequences at residues in and around this pocket indicate that it is present in many HLAs including HLA-B27 and HLA-B44. In HLA-B27, position 45 is a negatively charged glutamic acid, and in HLA-B44 it is a positively charged lysine. Thus the '45 pocket' is formally reminiscent of the specificity site in serine proteases, a pocket that is hydrophobic in chymotrypsin but negatively charged in trypsin. In HLA-peptide complexes it is possible that this pocket binds nonpolar side chains in HLA-Aw68, and positively and negatively charged polar side chains in HLA-B27 and HLA-B44, respectively. Large substitutions at position 67 at the entry of the pocket could close it in some allelic products such as HLA-B7 (Tyr at position 67). Most of the residues lining the pocket and around its opening are polymorphic, thereby possibly changing the pocket specificity in different HLA alleles. Taurog and El-Zaatari[22] have presented evidence that raises the possibility that a free cysteine at position 67 in HLA-B27 is covalently modified in the HLA-B27-linked disease ankylosing spondylitis. Given the location of position 67 at the edge of a negatively charged pocket in HLA-B27, it might be possible to design, perhaps from an antigenic peptide, a covalent blocking agent incorporating a positive charge to complement the '45 pocket' and a thiol to react with cysteine 67. Consideration of pockets might be sufficient to predict peptide binding even when alleles differ by a substantial number of residues. The sequence of HLA-Aw69, which is a natural hybrid of $\alpha_2$- and $\alpha_3$-domains from HLA-A2 and the $\alpha_1$-domain from HLA-Aw68 (ref. 3), indicates that the '45 pocket' would

be present but that the Asp 74 pocket would probably be blocked by Arg 97 and Tyr 116 (as in HLA-A2), possibly explaining why one HLA-A2-specific peptide antigen also interacts with HLA-Aw69 (ref. 23).

Comparison of the three-dimensional structures of HLA-Aw68 and HLA-A2 demonstrates how different pockets or subsites within the cleft of the class I histocompatibility glycoproteins are produced by substitutions at positions of amino-acid polymorphism (compare Fig. 3*c* and *d*). The difference between these two allele products is typical of MHC allelic diversity, and so the same degree of structural change in pockets and subsites should be observed for other alleles. In the case we have described, polymorphism results in limited local structural changes; the backbone structure is generally not altered (Fig. 1*b*), and residues identical in both structures remain essentially unperturbed[18]. The observed effects of polymorphism were on the shape (for example, intrusions and ridges) and charge-distribution of the antigen-binding site, and on some local pockets, resulting in changes in their specificity and in their distribution in the site. These structural changes provide the basis for allelic specificity in peptide affinity[11,15,16] and the resultant non-responsiveness (and possibly hyper-responsiveness) to certain immunological challenges[10]. Because amino-acid side chains from antigenic peptides could bind in these pockets, peptides binding to one MHC allelic product could then have a particular pattern of side chains[24,25] to fit the pockets. Moreover, because the affinity of antigenic peptides for individual histocompatibility glycoproteins may depend on only a few of an antigenic peptide's side-chains fitting into a subset of the available pockets[26,27], a broad range of peptides could be accommodated by the various combinations of filled and empty pockets. □

Received 16 August; accepted 12 October 1989


1. Bjorkman, P. J. *et al. Nature* **329,** 506–512 (1987).
2. Bjorkman, P. J. *et al. Nature* **329,** 512–518 (1987).
3. Holmes, N. & Parham, P. *EMBO J.* **4,** 2849–2854 (1985).
4. Koller, B. H. & Orr, H. T. *J. Immun.* **134,** 2727–2733 (1985).
5. Townsend, A. R. M., Gotch, F. M. and Davey, J. (1985) *Cell* **42,** 457 (1985).
6. Germain, R. M. *Nature* **322,** 687 (1986).
7. Bevan, M. *Nature* **325,** 192 (1987).
8. Townsend, A. & Bodmer, H. *A. Rev. Immun.* **7,** 601–624 (1989).
9. Parham, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **85,** 4005–4009 (1988).
10. Babbitt, B., Allen, P., Matsueda, E., Haber, E. & Unanue, E. *Nature* **317,** 359–361 (1985).
11. Buus, S., Sette, A., Colon, S. M., Miles, C & Grey, H. M. *Science* **235,** 1353–1358 (1987).
12. Sette, A., Buus, S., Colon, S. M., Miles, C. & Grey, H. M. *J. Immun.* **141,** 45–48 (1988).
13. Townsend, A. R. M. *et al., Cell* **44,** 959–968 (1986).
14. Maryanski, J. L., Pala, P., Curadin, J., Jordan, B. B. & Cerotini, J. C. *Nature* **324,** 578–579 (1986).
15. McDevitt, H. O. & Tyan, M. L. *J. expl. Med.* **128,** 1–11 (1968).
16. Fox, B. S., Carbone, F. R., Germain, R. N., Paterson, Y. & Schwartz, R. H. *Nature* **331,** 538–540 (1988).
17. Madden, D. R. *et al. Cold Spring Harb. Symp. Quant. Biol.* **54,** (In the press).
18. Garrett, T. P. J., Saper, M. A. & Wiley, D. C. *J. molec. Biol.* (submitted).
19. Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *J. molec. Biol.* **188,** 205–210 (1985).
20. Salter, R. D. *et al. Nature* **338,** 345–347 (1989).
21. Connolly, J. M., Potter, T. A., Wormstall, E-M. & Hansen, T. H. *J. exp. Med.* **168,** 325–341 (1988).
22. Taurog, J. D. & El-Zaatari, A. K. *J. clin. Invest.* **82,** 987–992 (1988).
23. Bodmer, H. C., Gotch, F. M. & McMichael, A. J. *Nature* **337,** 653–655 (1989).
24. Sette, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **86,** 3296–3330 (1989).
25. Rothbard, J. B. & Taylor, W. R. *EMBO J.* **7,** 93–100 (1988).
26. Jardetzky, T. S. *et al. EMBO J.* (submitted).
27. Maryanski, J. L., Verdini, A. S., Weber, P. C., Salemme, F. R. & Corradin, G. *Cell* (submitted).
28. Ellis, S. A., Strachan, T., Palmer, M. S. & McMichael, A. J. *J. Immun.* **142,** 3281–3285 (1989).
29. Connolly, M. L. *Quantum Chemistry Program Exchange Bull.* **1,** 74.
30. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235,** 458–460 (1987).
31. Brünger, A. T. in *Topics in Molecular Biology* (ed. Goodfellow, J. M.) (In the press).


ACKNOWLEDGEMENTS. We thank Anastasia Haykov for technical assistance, Ted Jardetzky and Peter Parham for comments on the manuscript, and acknowledge support from the NIH and Howard Hughes Medical Institute. We also thank Drs Dean Mann (NIH), D. Michael Strong and James Wood (US Naval Research Unit) and Don Giard (MIT Cell Culture Facility) for provision of cells. Coordinates of HLA-Aw68 have been deposited in the Brookhaven Protein Data Bank (Accession no. 2HLA).

Printed in Great Britain by Turnergraphic Limited, Basingstoke, Hampshire