Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 8269–8273, August 1995
Immunology

A general model of invariant chain association with class II major histocompatibility complex proteins

(antigen presentation/protein structure/molecular modeling)

CHRISTOPHER LEE AND HARDEN M. MCCONNELL

Department of Chemistry, MC 5080, Stanford University, Stanford, CA 94305

Contributed by Harden M. McConnell, June 1, 1995

ABSTRACT **The binding of invariant chain to major histocompatibility complex (MHC) proteins is an important step in processing of MHC class II proteins and in antigen presentation. The question of how invariant chain can bind to all MHC class II proteins is central to understanding these processes. We have employed molecular modeling to predict the structure of class II-associated invariant chain peptide (CLIP)–MHC protein complexes and to ask whether the predicted mode of association could be general across all MHC class II proteins. CLIP fits identically into the MHC class II alleles HLA-DR3, I-A^k, I-A^u, and I-A^d, with a consistent pattern of hydrogen bonds, contacts, and hydrophobic burial and without bad contacts. Our model predicts the burial of CLIP residues Met-91 and Met-99 in the deep P1 and P9 anchor pockets and other detailed interactions, which we have compared with available data. The predicted pattern of I-A allele-specific effects on CLIP binding is very similar to that observed experimentally by alanine-scanning mutations of CLIP. Together, these results indicate that CLIP may bind in a single, general way across products of MHC class II alleles.**

Invariant chain plays an important role in the processing of major histocompatibility complex (MHC) class II proteins, by blocking binding of antigenic peptides until appropriate. To do this , it has the unique capability of binding to the products of all MHC class II alleles, apparently overcoming the extensive polymorphism that gives each protein a unique motif of antigenic peptides that it binds and presents. The question of how invariant chain blocks antigen binding (suggesting a common binding site), yet is apparently unconstrained by known antigen-binding motifs, has remained difficult to resolve, with recently reported evidence both for and against its binding like a conventional antigen, within the groove of the MHC class II molecule (1–4).

We have sought to address these questions by using computer modeling to predict a detailed structure of invariant-chain peptide to MHC class II protein and to assess the properties of this hypothetical structure against available data. The unusual structural characteristics of MHC protein make it particularly suited to modeling: extensive polymorphism, within an open but tightly packed antigen binding site, framed by a highly conserved backbone fold. The striking conservation of the backbone structure of MHC proteins from allele to allele—e.g., in class I, 0.45 Å C^α rms between HLA-A2 and HLA-Aw68 (5)—provides an important foundation for accurate modeling of the structural differences between products of different alleles, the polymorphic side chains. The dense packing of the groove of MHC proteins tightly constrains the possible conformations of these side chains.

We have modeled the complex between the product of HLA-DR3 and the class II-associated invariant chain peptide (CLIP), an important test case for the accuracy of modeling, whose structure is being solved by x-ray crystallography. Furthermore, we have used modeling's ability to predict and compare the structures of multiple MHC alleles, to assess in detail the possibility that CLIP might be binding to all class II MHC proteins in an identical way. Our results provide several areas of new insights: (*i*) a detailed structural model of CLIP-MHC binding and interactions and predictions that can be used to test the model, (*ii*) evidence of a general mode of association of CLIP across all class II MHC proteins, and (*iii*) specific structural understanding of how CLIP can bind to the products of many different MHC alleles.

MATERIALS AND METHODS

Molecular Modeling. All modeling calculations were performed with self-consistent ensemble optimization (SCEO) as described (6, 7), by using the program CARA through the LOOK interface (Molecular Applications Group). We started from the DR1-HA crystal structure (8), Protein Data Bank code 1DLH (9). SCEO calculations were generated by using linear cooling from 6000 K to 298 K over 15 cycles with "heavy" data collection, followed by 10 cycles of equilibration at 298 K. Structure predictions for each residue were taken from the highest probability conformation of the residue in the calculated ensemble of the final equilibration cycle. All calculation steps and parameters were as described (7); however, the jump frequency was reduced to one jump per 100 steps, and the minimum conformational sample per cycle was increased to 5000.

To build the initial model of DR3, side-chain coordinates for all polymorphic positions (9, 10, 11, 12, 13, 26, 28, 30, 31, 32, 37, 47, 71, 73, 74, 77, 86, 96, 140, and 149) were deleted and modeled in one SCEO run, with no peptide coordinates. To build the DR3–CLIP model, the HA backbone coordinates from 1DLH were superimposed on our DR3 model, and CLIP residues 89–100 (SKMRMATPLLMQ) were constructed on this framework by one SCEO run in which all CLIP side chains were modeled *ab initio*, while the protein coordinates were held fixed. Subsequent to submission, a bad contact was observed between CLIP-94 Ala C^β and Arg-β74; therefore, an additional SCEO run to allow this side chain to adjust for the presence of peptide was performed, allowing β70, β74, and CLIP-92 to move. This slightly more refined model is not shown in this paper; coordinates are available from the authors. To model CLIP in I-A^u, I-A^k, and I-A^d, we superimposed the CLIP model coordinates upon our existing models of these alleles (10) and ran SCEO optimization on each of the complexes, allowing all CLIP side chains, as well as α- and β-chain peptide-binding domain polymorphic residues, to move.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MHC, major histocompatibility complex; CLIP, class II-associated invariant chain peptides; SCEO, self-consistent ensemble optimization.

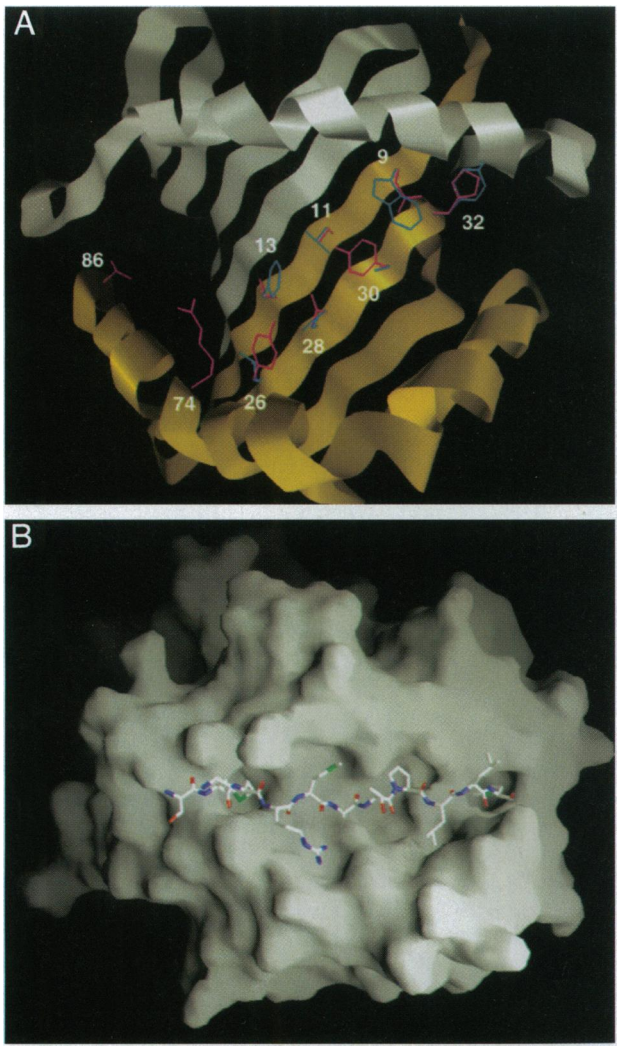

FIG. 1. Modeling of DR3–CLIP complex. (*A*) DR3 polymorphisms affecting the peptide-binding cleft. The α (white) and β chains (orange) of the DR3 model are shown in ribbon representation, with selected polymorphic side chains from DR3 (red) compared with side chains from DR1 (blue). The figure was generated by using LOOK (Molecular Applications Group, Palo Alto, CA). (*B*) Calculated molecular surface of the DR3 model with CLIP shown as skeleton bonds. The figure was generated by using GRASP (14).

RESULTS AND DISCUSSION

We have modeled the class II MHC allele DR3 by using a completely automated procedure that predicts the coordinates of polymorphic side chains upon a fixed backbone framework, as described (7, 10, 11). Tests of this procedure on the murine class I allele H-2K^b, starting from the structure of HLA-A2, gave a coordinate rms versus the crystal structure of 0.97 Å for the polymorphic residues (C.L. and H.M.M., unpublished results) and overall showed nearly the same level of similarity to the crystal structure as the two, independently solved crystal structures of H-2K^b (12, 13) share with each other. We used the crystal structure of HLA-DR1 (8) as the starting point for the model and modeled *ab initio* all 20 polymorphisms between the DR1 and DR3 β chains. Within the peptide-binding domain, the main focus of this study, the DR1 and DR3 β chains share about 82% identity.

Nine of the polymorphisms seem likely to have significant effects on peptide binding within the cleft (Fig. 1*A*). Several changes deepen the cleft along the interface between α and β chains (Trp-9 → Glu, Leu-11 → Ser, Phe-13 → Ser), while two adjacent changes create raised areas on the cleft floor (Leu-25 → Tyr, Cys-30 → Tyr). Glu-β9 sits at the base of the P6 and P9 pockets and could favor a basic anchor residue. Gly-86 → Val reduces one side of the P1 pocket, consistent with experimental observations that this mutation modulates P1-pocket specificity (15–17).

We have modeled CLIP within the cleft of HLA-DR3, using as a framework the backbone of the influenza virus hemagglutinin peptide cocrystallized with DR1 (8). Several criteria guided out placement of the CLIP-81–104 peptide: (*i*) some experimental evidence indicates that CLIP binds within the cleft much like antigenic peptides (2–4). (*ii*) The N-terminal residues 81–87 do not contain an appropriate binding motif. In particular, the high proline content (>50%) seems unlikely to provide a good fit to the cleft. (*iii*) We sought a conformation that would place a hydrophobic residue in the P1 anchor pocket and also in the P7 and P9 pockets. We considered seven different placements, three of which met this criterion. (*iv*) The existence of a proline at position 96 in CLIP led us to examine which locations in the cleft could accommodate its unusual conformation requirements. Most positions within the extended conformation of the peptide in the cleft fit the proline ϕ torsion reasonably well, but one of the three candidates (placing Pro-96 in the P4 position; $\phi = 130°$) did not. Of the remaining two candidates, one placed Pro-96 in a pocket (P6) where it fit well, while the other created strong steric clashes to the proline side chain (P9). We chose the former. (*v*) This alignment is consistent with CLIP alanine-scanning data on binding to DR1 and DR17 (3). Using this alignment, we modeled all residues of CLIP within our DR3 structure. The CLIP–DR3 complex model was deposited as a blind prediction with several independent investigators in February 1995 and submitted to the upcoming Asilomar modeling evaluation conference.

The model predicts Met-91 in the deep P1 anchor pocket, Met-93 in the P3 pocket, Pro-96 in P6, Leu-97 in P7, and Met-99 in P9 (Fig. 1*B*). Lys-90, Arg-92, Thr-95, and Leu-98 are predicted to point upwards into solvent. Our model indicates that CLIP can fit directly into the cleft like an antigenic peptide, without bad contacts, and predicts a variety of detailed interactions between DR3 and CLIP (Fig. 2). Our model gives a good fit of CLIP into the DR3 peptide-binding cleft by several criteria: pattern of hydrophobic burial of peptide side chains, pattern of hydrogen bonding to the MHC protein, and absence of conformational strains or bad contacts. One strained interaction was observed, between CLIP Pro-96, and DR3 Asn-α62, which normally hydrogen bonds to the peptide amide group of the P6 peptide residue. The placement of Pro-96 at the P6 position makes this hydrogen bond impossible. It is striking that the model fills the two deeply buried anchor pockets (P1 and P9) with methionine, a rare amino acid. Methionine has the unique property of providing a large hydrophobic side chain with torsional freedom for every atom in the side chain. This unmatched flexibility may allow CLIP to satisfy the requirement for hydrophobic burial in these anchor pockets in some allelic products—e.g., P1 in DR17—and still fit to other allelic products in which these pockets are modified or reduced—e.g., by the β86 mutation in DR3.

To investigate the hypothesis that CLIP might bind in a general way across all class II MHC proteins, we have modeled CLIP with the murine class II I-A^k, I-A^u, and I-A^d, whose structures have been modeled by these methods (10). We placed CLIP in the peptide-binding clefts of these proteins in an identical registry to that in DR3, and modeled all side chains of the peptide *ab initio*. This binding model fit CLIP into all three proteins (Fig. 3) without bad contacts or repulsive interactions. In contrast, the hemagglutinin-306–318 peptide placed by the same procedure into the I-A^k allele had steric clashes in both the P1 (Tyr-308) and the P9

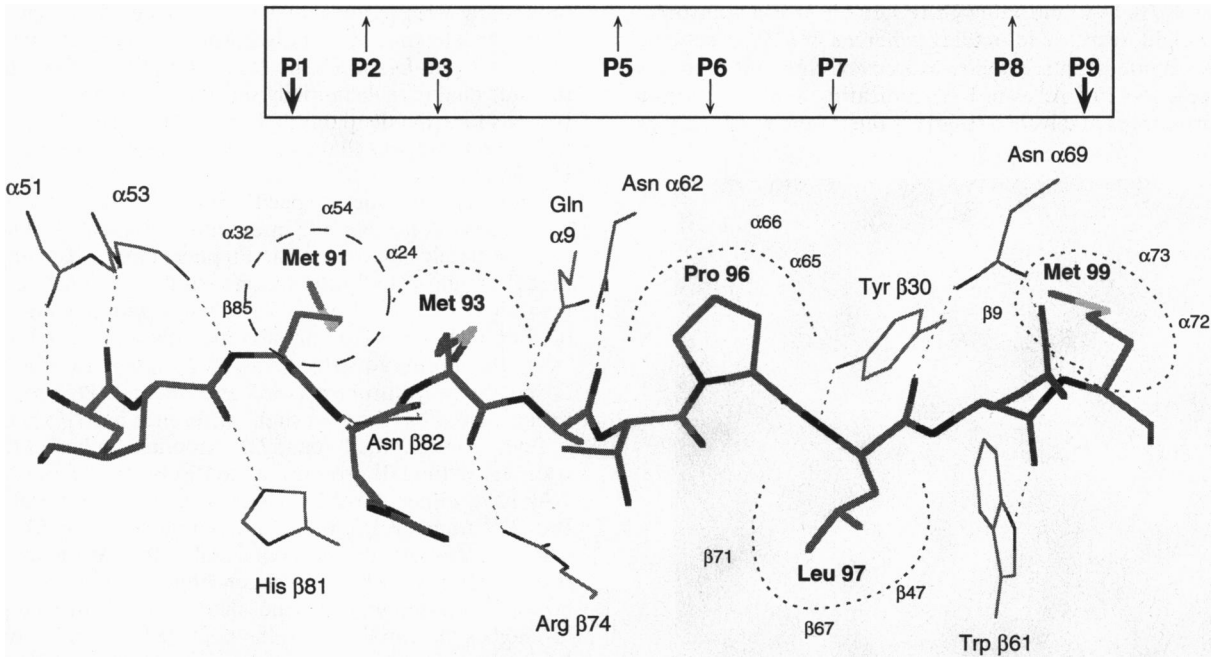

FIG. 2. Interactions between CLIP and the DR3 peptide-binding cleft. CLIP is shown in skeleton bonds (thick lines, labeled in bold); DR3 residues (thin lines) making hydrogen bonds (straight, dotted lines) to CLIP are also shown. The five DR3 pockets that bind CLIP side chains are indicated with curved, dashed lines: the deeply buried anchor pockets P1 and P9 and the P3, P6, and P7 pockets. DR3 residues contacting CLIP in these pockets are marked by small labels. The schematic above the figure presents the pattern of hydrophobic burial (downward arrows) versus accessibility (upward arrows) for the side chains of CLIP. The figure was generated by using LOOK (Molecular Applications Group).

(Leu-316) pockets (data not shown). The I-A–CLIP complexes showed similar patterns of hydrogen bonds, contacts, and hydrophobic burial as the DR3–CLIP complex.

These I-A–CLIP models suggest detailed predictions of allele-specific effects on CLIP binding to these class II proteins which can be compared with experimental data. Among I-A^k, I-A^u, and I-A^d, there are about 17 amino acid differences in areas of the cleft that could affect peptide binding. Mapping allele-specific polymorphisms—i.e., residues where one allele differs from the amino acid shared by the other two alleles—to the molecular surfaces of these complexes (Fig. 3) predicts interactions with individual CLIP side chains that should be detected by CLIP alanine-scanning mutagenesis experiments (1). By scrutinizing such data for allele-specific effects—i.e., a CLIP mutation that affects binding to two of the alleles identically, but differently to the third—interactions between individual allele-specific polymorphic residues in MHC proteins and individual residues in CLIP can be identified.

Trp-β61 (DR numbering, which will be used throughout this paper) is a highly conserved residue across both class II and class I MHC proteins, one of five residues forming the P7 pocket in DR1 (8). I-A^u contains an unusual polymorphism in which Trp-β61 is mutated to a Tyr, shifting the wall of the P7 pocket in our model (10) and possibly favoring a smaller P7 side chain. This result is supported by experimental data on the effects on mutating CLIP residues to alanine, which were measured by changes in the ability of CLIP to compete with standard peptides for binding to I-A^k, I-A^u, and I-A^d (1).

Just 4 of the 18 positions in CLIP that were tested showed strong allele-specific binding effects: Met-91, Met-93, Pro-96,

Table 1. Predictions of MHC protein–CLIP association model

CLIP mutant	Model prediction	Observed effect*		
		U	K	D
L97A	β61 polymorphism should modulate specificity in the P7 pocket. K and D share conserved Trp; U has unusual Tyr mutation, predicting a U-specific binding effect at this site. Further tests: kku61 should respond to L97A like U; uuk38,61 should respond to L97A like K.	+++	−	−
P96A	β9, the only K-specific polymorphism in the cleft, should effect P6 specificity, possibly through shifting the conserved Tyr-β30 residue. Further tests: kku9 should respond to P96A like U; uuk9 should respond to P96A like K.	0	++	+
M99A	Model predicts no polymorphisms distinguishing U, K, and D within the P9 pocket. Further tests: kku38,61 could have a slight effect on M99A binding.	+++	+++	+++
M91A	Model predicts K similar to human DRs within the P1 pocket, favoring a larger hydrophobic amino acid. U changes hydrogen bonding to P1 residue backbone by mutation at β81. D reduces size of the P1 pocket via Pro-β86. Further tests: T86S and T86L effects on CLIP binding should not be observed with CLIP M91A.	++	−−−	+

See ref. 10 for a description of the kku and uuk mutant I-A^k, I-A^u proteins.
*+, <20% increase; ++, <50% increase; +++ >80% increase; 0, no effect; −, <20% decrease; −−−, >80% decrease.

and Leu-97. In I-A^k, mutating CLIP Leu-97 → Ala appears to cause a slight decrease in binding, whereas in I-A^u, it actually improves binding substantially. Moreover, this mutation has the same effect in I-A^d as in I-A^k, indicating that a U-specific polymorphism is involved. Trp-β61 is one of the few locations interacting with peptide that has a U-specific polymorphism (Fig. 3*A*). The only other polymorphism among the P7 contact residues is the D-specific mutation Ile-β28 → Thr. However, this side chain is substantially shorter than in DR1 (Glu), and, from its location deep on the floor of the β sheet, is less likely to have interactions that can distinguish the P7 side chain (see Fig. 3*C*).

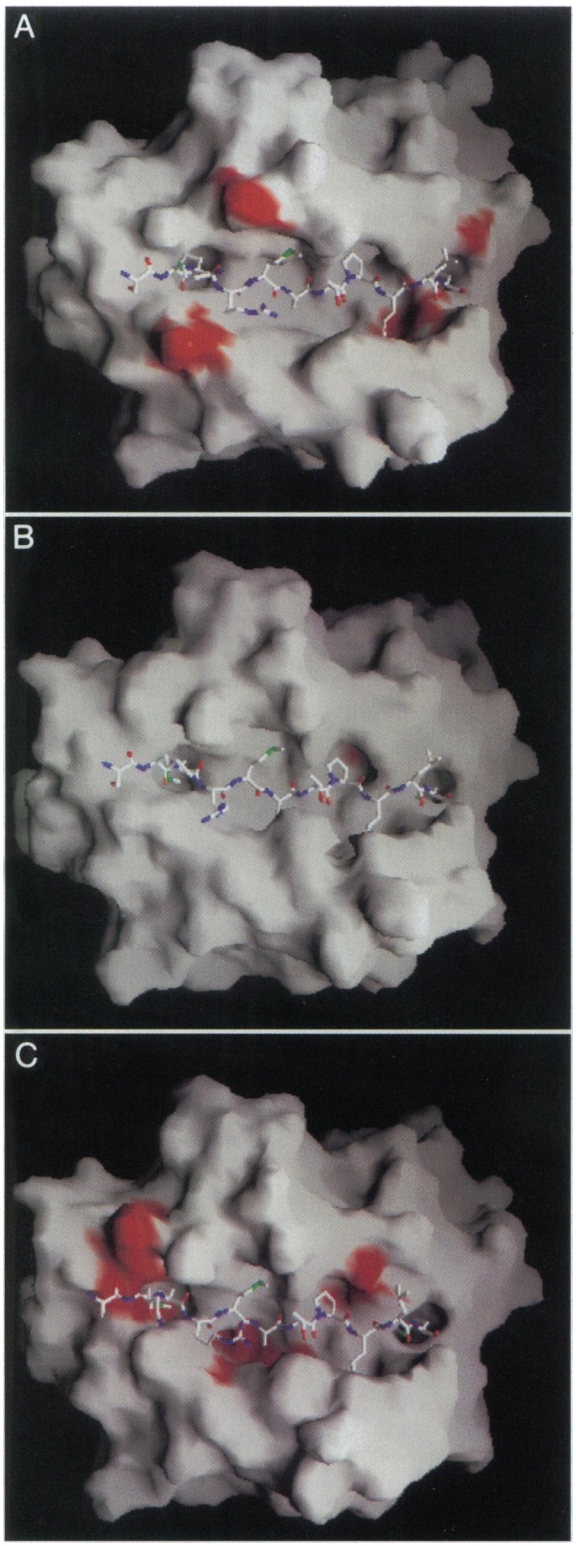

FIG. 3. Proposed CLIP–I-A complex structures. Complexes were modeled by using I-A^u (*A*), I-A^k (*B*), or I-A^d (*C*). Allele-specific polymorphisms within the peptide-binding cleft are colored red on the molecular surface of each binding protein. The figure was generated using GRASP (14).

There is only one K-specific polymorphism that appears liable to influence peptide interactions: Val-β9 → His (10). It is immediately adjacent to the P6 pocket and according to the model should affect interactions with CLIP Pro-96. Experimentally, mutating Pro-96 → Ala had essentially no effect on binding to I-A^u or I-A^d but significantly improved binding to I-A^k; the observed effect was K specific. Two additional, D-specific polymorphisms, α65 and α66, are P6 contacts; the experimental data show a slight difference in I-A^d vs. I-A^u (1).

Comparison of the I-A and DR structures indicates that I-A^k is similar to the DR proteins in the P1 pocket but that I-A^u and I-A^d have important differences. I-A^d has the mutation β86 Pro (DR numbering), which blocks one side of the P1 pocket and may cause further backbone shifts. In I-A^u, the conserved residue His-β81, which hydrogen bonds to the CLIP-90 carbonyl, is mutated to Tyr, either deleting the hydrogen bond or forcing a substantial shift in the peptide backbone around the P1 pocket. Experimentally, changing CLIP Met-91 → Ala causes a drastic decrease in binding to I-A^k, indicating a requirement for a large hydrophobic anchor at P1, as in DR1. In I-A^u, and to a lesser extent in I-A^d, the smaller alanine side chain is preferred. Our model can be further tested by assaying binding of CLIP Met-91 → A to existing I-A^u β86 → Leu, Ser mutants. These mutants have large effects on binding of regular CLIP (1); our model predicts that CLIP Met-91 → A should be insensitive.

The only allele-specific polymorphisms near the P3 pocket are D specific (α11 and α28); experimentally, mutating CLIP Met-93 → Ala indeed gives a similar response in U and K, but different in D. At P9, only the U-specific mutation Val-β38 → Leu, deep in the cleft, seems likely to affect side chain specificity. The nearby D-specific mutation Ile-α72 → Val removes a single methyl group, which contacts only the C^β atom of CLIP Met-99, and thus would be predicted not to distinguish between methionine and alanine at the P9 position. Experimentally, mutating CLIP Met-99 → Ala causes a roughly equivalent improvement in binding to all three proteins.

Table 1 summarizes these results and predictions. Broadly interpreted, they suggest that CLIP not only can bind within the cleft of the MHC protein in the same manner as an antigenic peptide but also that it binds in an identical way to products of various class II alleles, at least for the human DRs and murine I-As. We have identified structural characteristics of CLIP which in our models appear to be important for binding the products of multiple alleles (Table 2). These features are not proposed to be optimal for binding to any given protein, but rather to permit CLIP to fit all allelic products. Further experiments can directly test our predictions. Mutants of I-A^k and I-A^u have recently been constructed which switch residues β9, β38, and β61 to the amino acid of the opposite allele, in the contexts of both I-A^k and I-A^u (10). Our models predict that these polymorphisms modulate the specificity of class II interactions with CLIP residues 96 and 97, respectively, and possibly 99 (Table 1). By assaying binding of the appropriate alanine-substituted CLIP peptides on these mutants, the predicted CLIP–MHC protein interactions can be specifically tested.

We wish to thank Drs. C. Beeson, P. Cresswell, M. Davis, M. Liang, and P. Parham for their advice and critical reading of this manuscript. This work was supported by the Office of Naval Research (Grant

Table 2. CLIP characteristics important for global binding to MHC proteins

Residue	Comments/analysis
Met-91	Rare, large hydrophobic residue with torsional freedom for every side-chain atom; allows binding to proteins that require a large hydrophobic residue at P1—e.g., DR17—but maximal flexibility sufficient to also bind those with an altered or reduced P1 pocket
Met-93	Hydrophobicity and flexibility (like Met-91) in P3 pocket
Ala-94	Smallest nonglycine side chain, gives hydrophobic contact, and avoids steric clash with surrounding polymorphisms in P4 pocket
Pro-96	Medium hydrophobic residue in P6 pocket; unusually stiff, may help enforce consistent mode of binding to different MHC proteins
Leu-97	Large hydrophobic residue in P7 pocket, appears likely to fit this position in all proteins
Met-99	Hydrophobicity and flexibility (like Met-91) in P9 pocket

ONR-N00014-90-J-1407). C.L. is a postdoctoral fellow of the American Cancer Society (Grant PF-4220).

1. Gautam, A. M., Pearson, C., Quinn, V., McDevitt, H. O. & Milburn, P. J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 335–339.
2. Liang, M. N., Beeson, C., Mason, K. & McConnell, H. M. (1995) *Int. Immunol.*, in press.
3. Malcherek, G., Gnau, V., Jung, G., Rammensee, H. & Melms, A. (1995) *J. Exp. Med.* **181,** 527–536.
4. Sette, A., Southwood, S., Miller, J. & Appella, E. (1995) *J. Exp. Med.* **181,** 677–683.
5. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. (1989) *Nature (London)* **342,** 692–696.
6. Lee, C. & Levitt, M. (1991) *Nature (London)* **352,** 448–451.
7. Lee, C. (1994) *J. Mol. Biol.* **236,** 918–939.
8. Stern, L. J., Brown, J. H., Jardetzky, T. J., Gorga, J. C., Urban, R. G., Strominger, J. L. & Wiley, D. C. (1994) *Nature (London)* **368,** 215–221.
9. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542.
10. Tate, K. M., Lee, C., Edelman, S., Carswell-Crumpton, C., Liblau, R. & Jones, P. P. (1995) *Int. Immunol.* **7,** 747–761.
11. Lee, C. & Subbiah, S. (1991) *J. Mol. Biol.* **217,** 373–388.
12. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257,** 919–927.
13. Zhang, W., Young, A. C. M., Imarai, M., Nathenson, S. G. & Sacchettini, J. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8403–8407.
14. Nicholls, A., Sharp, K. & Honig, B. (1991) *Proteins* **11,** 281–296.
15. Busch, R., Hill, C. M., Hayball, J. D., Lamb, J. R. & Rothbard, J. B. (1991) *J. Immunol.* **147,** 1292–1298.
16. Demotz, S., Barbey, C., Corradin, G., Amoroso, A. & Lanzavecchia, A. (1993) *Eur. J. Immunol.* **23,** 425–432.
17. Newton-Nash, D. K. & Eckels, D. D. (1993) *J. Immunol.* **150,** 1813–1821.