The three-dimensional structure of a class I major histocompatibility complex molecule missing the α3 domain of the heavy chain

EDWARD J. COLLINS, DAVID N. GARBOCZI, MICHAEL N. KARPUSAS*, AND DON C. WILEY†

Department of Molecular and Cellular Biology, Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138

ABSTRACT   Class I major histocompatibility complex (MHC) molecules are ternary complexes of the soluble serum protein β_2-microglobulin, MHC heavy chain, and bound peptide. The first two domains ($\alpha1$, $\alpha2$) of the heavy chain create the peptide binding cleft and the surface that contacts the T-cell receptor. The third domain ($\alpha3$) associates with the T-cell co-receptor, CD8, during T-cell recognition. Here we describe the x-ray crystal structure of a human class I MHC molecule, HLA-Aw68, from which the $\alpha3$ domain has been proteolytically removed. The resulting molecule shows no gross morphological changes compared to the intact protein. A decameric peptide complexed with the intact HLA-Aw68 is seen to bind to the proteolized molecule in the conventional manner, demonstrating that the $\alpha3$ domain is not required for the structural integrity of the molecule or for peptide binding.

Class I major histocompatibility complex (MHC) molecules present cytoplasmic peptides to cytotoxic T cells as part of a mechanism of immunological surveillance (reviewed in ref. 1). The extracellular portion of the heavy chain consists of three domains that are each involved in T-cell recognition. The $\alpha1/\alpha2$ domains form the peptide binding cleft. The $\alpha3$ domain interacts with the T-cell co-receptor CD8 (2–5) in the majority of class I MHC molecules, and CD8 may also interact with the $\alpha2$ domain (6). β_2-Microglobulin (β_2m) has a role in enhancing peptide binding in vitro (7, 8) yet is not absolutely required for the presentation of viral epitopes in vivo (9).

Although the $\alpha3$ domain appears to be required for interaction with CD8, peptide binding is retained when the entire $\alpha3$ domain is removed by exon deletion (10). Here we describe the crystallographic structure of the human class I MHC molecule HLA-Aw68 complexed in vitro (11) with a decameric peptide and truncated proteolytically to delete the $\alpha3$ domain.‡ Despite deletion of $\alpha3$ the molecule is structurally unchanged and peptide is bound conventionally.

MATERIALS AND METHODS

Peptide. A peptide, EVAPPEYHRK, previously eluted from human HLA-Aw68 (12) was synthesized on an Applied Biosystems peptide synthesizer (model 431A) and purified by C_{18} reversed-phase HPLC (Waters). The identity of the purified peptide was confirmed by fast atom bombardment mass spectrometry (Harvard Chemistry Department.)

Preparation of EVAPPEYHRK/HLA-Aw68 Cocrystals. Total cellular mRNA was isolated from the human lymphoblastoid cell line LB and cDNA was produced by standard protocols. Oligonucleotide primers, containing EcoRI and HindIII restriction sites were used to modify an initial clone that contained a mutation and to generate a sequence for residues 1–274 of the HLA-Aw68 heavy chain. The 5′ primer

contained a ribosome binding site, a start codon, and a coding sequence to repair the mutation made in the amplification of the cDNA. Plasmid template was amplified by 15 cycles of polymerase chain reaction, isolated with glassmilk (Geneclean), gel purified, digested with EcoRI/HindIII. The DNA was cloned into the plasmid pHN1+ (13), DNA was sequenced entirely, and insoluble protein ($\approx$20 mg/liter) was produced. Heavy chain was folded in the presence of β_2m and peptide as described (11, 14). Production of HLA-Aw68, as monitored by HPLC gel filtration (11), was peptide dependent. After 36 hr at 10°C, the protein was concentrated and purified by HPLC gel filtration (11). Purified protein was concentrated (20 mg/ml in 25 mM Mes, pH 6.5) by ultrafiltration (Amicon Centriprep-10) and crystallized by vapor diffusion (room temperature). Equal volumes of protein and well solution were mixed and suspended over PEG 8000 (10–20%)/25 mM Mes, pH 6.5. Crystals formed only in one drop that had been seeded with crystals of HLA-Aw68 purified from human cells and complexed with an influenza virus peptide (15, 16).

Proteolysis that removed the $\alpha3$ domain occurred adventitiously during crystallization. Immunoblotting an SDS/polyacrylamide gel of a washed crystal with a polyclonal antibody revealed a 23-kDa fragment corresponding to $\alpha1/\alpha2$. Subsequent incubations of intact HLA-Aw68/peptide complexes with proteases (1:500 molar ratio, protease/peptide, 37°C, 125 min; papain, trypsin, chymotrypsin, elastase, and subtilisin) indicated that a major product was the size expected for $\alpha1/\alpha2$ (data not shown).

X-Ray Data Collection and Space Group Determination. Crystals were transferred to increasing concentrations of glycerol (5%, 10%, and 20%) in 14% PEG 8000 and 25 mM Mes (pH 6.5) for 60 min at each step, were mounted in loops (diameter, $\approx$300 μm) of dental floss fiber (20 μm), and were flash-cooled to -165°C (17). Oscillation data (1.0°, 90 sec, 190 frames) were collected from a single crystal ($\approx$50 × 20 × 10 μm) at 235 mm crystal-to-detector distance using Fuji PhosphorImaging plates at the Cornell High Energy Synchrotron Source F1 beam line (wavelength, 0.908 Å). Diffraction intensities were indexed and integrated with DENZO (Z. Otwinkowski, Southwestern Medical Center, Dallas) and scaled as space group P1 [CCP4 programs ROTOVATA/AGROVATA and POSTREFINE (18)]. Visual inspection of pseudo-precession photographs generated with PLOTHKL (§) suggested MM symmetry and systematic absences along each axis indicating space group $P2_12_12_1$ ($a = 48.23$, $b = 58.05$, $c = 109.53$ Å). This was confirmed by comparing R_{merge} values calculated

---

---

across predicted mirror planes (12–15%) with nonmirror planes (61–74%). The data, which were 99% complete to 2.30 Å, were reintegrated and scaled (R_{merge} = 0.07 for data from 20.0 to 2.30 Å, 0.228 for the 2.39- to 2.3-Å shell).

Molecular Replacement. Rotation function calculations were performed with various search models based on HLA-Aw68 (12) (Table 1). A variety of resolution shells within the range of 10–3.0 Å and radii of integration were used. All searches were done using AMoRe (19) and MERLOT (20).

Crystallographic Refinement. The model determined by molecular replacement was refined as a rigid body using X-PLOR (21) and refined by iterative cycles of manual rebuilding and least-squares refinement using O (22) and X-PLOR. A typical refinement cycle consisted of 50 steps of positional refinement, a slow-cool procedure from 2000 K, 50 steps of positional refinement, and 15 steps of neighbor-restrained individual B refinement. Only refinement strategies that resulted in decreases in R_{free} (23) were used. A peptide model was added after six cycles when the entire length of the peptide could be traced and the R_{free} value was <35%. Ninety-eight waters were added in peaks >3σ within 2.5–3.4 Å of two hydrogen bond donor/acceptors and were retained in the model if during refinement the individual temperature factor stayed below 40 Å^2.

To minimize model bias an "all-omit" map was calculated (as described in ref. 24, except modified to use X-PLOR) using 10 slabs along Z and used in model building. The final model has 92% of its side chains in interpretable electron density. There are only two breaks in main-chain electron density, both in distal loops on β_2m. Real space fits of main-chain and side-chain atoms are 87% and 81% (22). Only 2% of all possible 9 residue windows score above 6.0 in ERRAT (25) and no windows of 21 residues are below 0.0 in three-dimensional profile (26). These analyses show no significant differences from intact HLA-Aw68. The final model has R_{free} = 30.4% for data >3σ (31.9% for all data; 1188 reflections) and an R_{work} of 20.7% (22.7% for all data; 12,084 reflections) (rms bonds = 0.008 Å, angles 1.916°) for 2245 protein and 98 solvent atoms. No residues are in disallowed regions in the Ramachandran plot and 91.6% are in the most favored regions. $R_{free} = (\Sigma_h |F_o - F_c|)/(\Sigma_h F_o), \forall h \in \{$free set$\}$. $R_{work} = (\Sigma_h |F_o - F_c|)/(\Sigma_h F_o), \forall h \in \{$working set$\}$. $R_{merge} = (\Sigma_{hkl} |I - \langle I \rangle|)/(\Sigma_{hkl} \langle I \rangle), \forall hkl \in \{$independent Miller indices$\}$.

RESULTS

The X-Ray Structure Reveals α3 Domain Removed. The first indication that the intact molecule had not crystallized was the observation that the volume of the asymmetric unit in the crystals of intact HLA-Aw68 previously studied was 130,000 Å^3 while the HLA-Aw68 peptide complex studied here was only 69% of that, 76,600 Å^3. Molecular replacement

calculations then indicated that the α3 domain was missing (Table 1).

Using the coordinates of intact HLA-Aw68 (12), the highest peak in rotation function and translation function calculations suggested a model with a correlation coefficient of 37.5% and a R_{factor} of 48.5% after rigid body refinement. However, visual inspection of a model of the crystal lattice indicated that the α3 domain overlapped almost completely with another molecule in the crystal. Further calculations with different resolution ranges and different search radii failed to resolve this packing problem. Calculations with subsets of the first model that included the α3 domain showed the same overlap, but a model with α3 deleted (α1/α2/β_2m) had a greatly improved correlation coefficient of nearly 60% and a R_{factor} below 40% (Table 1). This α3-deleted model packed in the lattice without overlaps and with apparent lattice contacts sufficient to form a crystal. Omit maps generated by omitting portions of the α1/α2/β_2m model showed clear electron density for the omitted portions and, in addition, contiguous electron density was seen in the peptide binding groove. No electron density was visible anywhere in the crystal for the α3 domain and the high quality of the density in the visible domains appeared to preclude the possibility that the apparent packing overlap may have resulted from a packing disorder or error in space group determination. The structure was refined to 2.3 Å and is comparable to other well-refined structures when examined for real space fit, ERRAT (25), and 3-D PROFILE (26) (see *Materials and Methods*). Biochemical analysis of the small crystal used in data collection confirmed that the α3 domain was not present. An immunoblot of the washed crystal revealed a weak band of apparent M_r 23,000 (α1/α2) and no band at M_r 35,000 (α1/α2/α3). Subsequently, limited proteolysis experiments performed with papain, subtilisin, trypsin, chymotrypsin, and elastase demonstrated a cleavage product of the approximate size of the α1/α2 domain.

Bound Peptide Structure. A "simulated annealing omit map" (27) shows unambiguous density for the peptide except for side chains P7 (Tyr) and P9 (Arg), which extend into solvent and may be disordered (Fig. 1). The peptide has 15 main-chain hydrogen bonds with the class I MHC molecule. These include 6 hydrogen bonds directly from conserved MHC residues to the termini of the peptide and one bridging through water (28, 29) (Fig. 2, underlined). Of the 21 main-chain atoms that have the potential for hydrogen bonding, 15 participate in hydrogen bonds to the MHC molecule directly or through water. This is similar to the extensive hydrogen bonding reported between peptides and other MHC molecules (15, 28, 30–32). In this peptide, individual side chains contact HLA-Aw68 at the anchor residues P2 and P9 (12, 28) and as secondary anchors (33, 34) at P1, P6, and P8. The side chain of P1 (Glu) points out of the peptide binding site, salt bridges with Arg-62, hydrogen bonds to Asn-63, and via a bound water

Table 1. Molecular replacement searches

Search model	α	β	γ	FA	FB	FC	C.C.	R_{factor}
1. α1/α2/α3/β_2m	35.2	22.1	281.6	0.12	0.34	0.25	37.5	48.5
2. α1/α2/α3	35.2	23.6	281.9	0.14	0.38	0.38	37.5	48.4
3. α1/α2/β_2m	34.0	21.8	282.4	0.12	0.26	0.29	58.2	39.7
4. α1/α2	35.3	23.0	280.1	0.16	0.25	0.38	50.9	43.2

Domain compositions of four search models are listed with the final rotation, translation, and rigid body fit parameters from the searches with data between 8.0 and 3.0 Å and a radius of integration from 4 to 25 Å. The next most significant peak was <65% of the height of the peaks listed. Model 1 is the full extracellular portion of HLA-Aw68. Model 2 is the heavy chain of the full HLA-Aw68 (residues 1–270). Model 3 is the first two domains of heavy chain (residues 1–183) and β_2m (residues 1–99). Model 4 is the first two domains of heavy chain (residues 1–183). Solutions to searches using β_2m alone are not listed, since none of the peaks was above the noise level. $R_{factor} = (\Sigma_h |F_o - F_c|)/(\Sigma_h F_o)$. C.C. = $[(\Sigma_h F_o F_c - \Sigma_h F_o \Sigma_h F_c)/N]/[\Sigma_h F_o^2 - (\Sigma_h F_o)^2/N]^{1/2}[\Sigma_h F_c^2 - (\Sigma_h F_c)^2/N]^{1/2}$. α, β, and γ are Eulerian rotation angles. FA, FB, and FC are fractional translations of unit cell edges a, b, and c, respectively. Differences in translations are primarily due to differences in the centers of mass of the different search models.
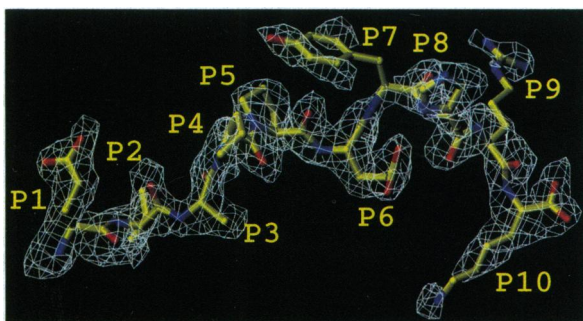
FIG. 1. Simulated annealing omit electron density map (27) for bound peptide EVAPPEYHRK. Peptide atoms and all protein or solvent atoms within 5 Å of the peptide were omitted before simulated annealing.

to Asn-66. The anchor residue, P2 (Val), fits into the P2 binding pocket (12, 35). P6 (Glu) is directed down and toward the α2 domain α-helix where it hydrogen bonds to Trp-156 and to the Nε of P8 (His). The side chain of P8 (His) is directed toward the α2 domain α-helix and may form a long hydrogen bond to the carbonyl oxygen of Ala-150. P9 (Arg) projects into solvent but in the crystal its position is stabilized by a salt bridge to Asp-98 of β_2m of another molecule in the crystal. The anchor residue, P10 (Lys), angles into the P9 binding pocket, forming hydrogen bonds with Asp-116 and through water to Asp-74 and Asp-77.

DISCUSSION

X-ray crystallography revealed that a class I MHC molecule bound to a decameric peptide had been proteolytically cleaved unexpectedly before crystallization to delete the α3 domain. The loop connecting α2 and α3 is susceptible to protease (36) and our subsequent proteolytic experiments indicate that HLA-Aw68 produced in *Escherichia coli* is readily cleaved to release α3. The most striking observation from the x-ray structure determination is the lack of structural changes that accompany the loss of the α3 domain (Fig. 3). There are only two residues with main-chain rms movements of >1.5 Å (Gly-16 in α1/α2, 2.2 Å; Asn-17 in β_2m, 1.59 Å). The absence of significant structural changes probably reflects the fact that the α3 domain has little contact with the α1/α2 domains [283.9-Å^2 solvent-accessible surface buried (12)] in the intact protein. The α3 domain interacts more with β_2m (587.3-Å^2 solvent-accessible surface) but those interactions are generally polar (37). In fact, free β_2m is very soluble so that it is reasonable to find its surface exposed to solvent in the α3-deleted MHC molecule. A small difference in the relative position of β_2m was found in the α3-deleted molecule when it was superimposed on the intact HLA-Aw68 using C$^\alpha$ coordinates in the α1/α2 domains. The first α1/α2 domains super-

FIG. 3. Domain structure of class I histocompatibility antigens. Region in dashed box has been proteolytically removed in the structure described here. The remainder of protein shows no gross morphological differences when compared to intact class I molecules and binds peptide normally.

impose with very good agreement (overall rms deviation, 0.51 Å), but β_2m shows a small rotation (2.17°) away from the position where α3 would be. At the most distal portion of β_2m this movement produces a 2-Å difference in atomic positions. This rotation reduces the surface area covered by β_2m on the α1/α2 domains by 41 Å^2.

Individual differences in amino acid positions are few and most appear to result from contact with other molecules in the crystal. For example Gly-16 has moved 2.2 Å, but Arg-17 and Gly-18 are in contact with other molecules in the crystals. Similarly, Tyr-10 of β_2m has moved but is stabilized by a crystal contact. Part of the α2 domain α-helical region (Ala-140 to Ala-149) is displaced laterally by 1.0 Å when compared to the structure with a mixture of peptides (12), probably as a result of crystal packing forces (see discussion in ref. 31) or to accommodate the longer peptide. The additional length of the decameric peptide when compared to a nonamer NP peptide bound to HLA-Aw68 (16) (Fig. 4A) is accommodated by a small lateral zigzag at P4 (Pro) and P5 (Pro) and a large vertical bulge of one residue (P7 Tyr). With the exception of these deviations, the overall path of the two peptides is very similar.
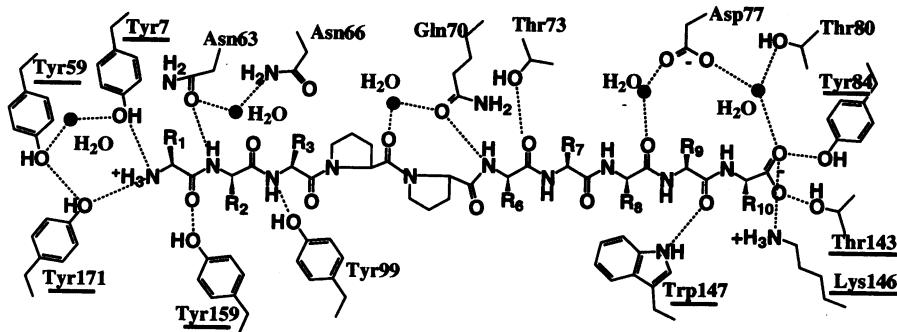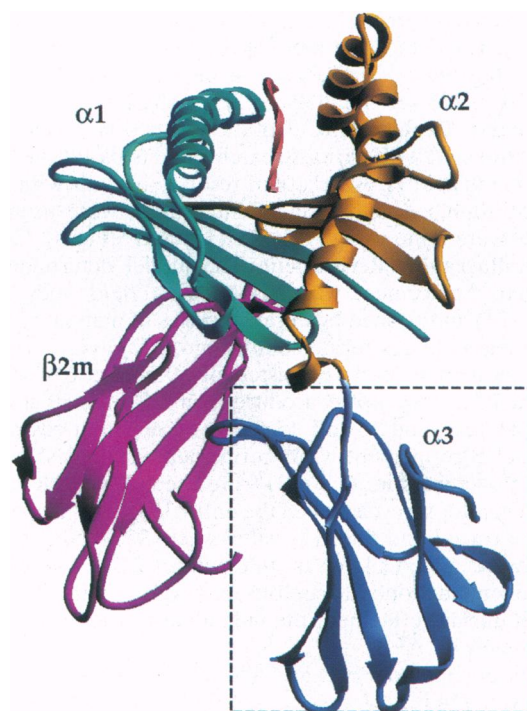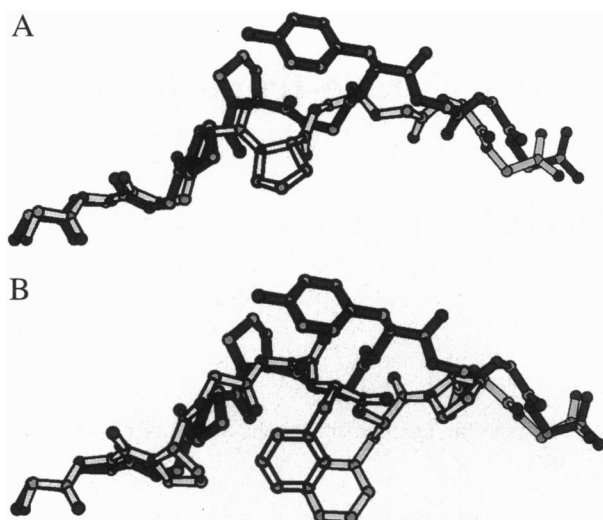

FIG. 2. Main-chain hydrogen bonds from peptide EVAPPEYHRK to HLA-Aw68. Residues conserved in either murine or human allotypes are underlined.

A

B

FIG. 4. (*A*) Overlay of EVAPPEYHRK decamer peptide with a nonamer bound to HLA-Aw68 (NP; KTGGPIYKR) (16). (*B*) Overlay of EVAPPEYHRK decamer peptide with a decamer bound to HLA-A2 (HP; FLPSDFFPSV) (31). Most side chains have been removed for clarity. Peptide EVAPPEYHRK is shaded black. Figure was generated by using MOLSCRIPT (38).

When the peptide binding domains are superimposed, the C^α positions of P1–P4 are 0.39–0.63 Å apart. They diverge considerably at P5 (3.76 Å) and reunite at P6 (1.26 Å). P7 (Tyr) bulges out of the site such that P8, P9, and P10 of the decamer align with P7, P8, and P9 of the monomeric NP peptide on HLA-Aw68. The peptide carboxylate of the decamer is shifted by 1 Å relative to the NP nonamer as previously observed in the other class I/decamer peptide structure, HLA-A2/hepatitis virus B peptide (31).

A comparison of the EVAPPEYHRK decamer to the hepatitis decamer (HP) bound to HLA-A2 (31) peptide shows striking differences (Fig. 4*B*). Both termini of both peptides are within 0.5 Å of one another when the peptide binding domains of the two molecules are superimposed; however, the additional lengths of the two peptides are accommodated by different paths, which appear to be allele specific. The HP decamer zigzags laterally across the HLA-A2 peptide binding site, projecting two hydrophobic phenylalanines toward the floor of the site. This is not possible in HLA-Aw68 because the HLA-Aw68 polymorphic residue Trp-156 fills the space. In contrast, the EVAPPEYHRK peptide zigzags vertically in and out of the site.

Overall, the loss of the α3 domain has had no apparent effect on how a peptide is bound. The EVAPPEYHRK peptide binds in a manner that is entirely consistent with that observed for a collection of endogenous peptides on HLA-Aw68 (12) and for HLA-A2 complexed to a decameric peptide (31). The observation of a stable structure of the α3-deleted MHC molecule implies that soluble class I HLA complexes could be produced with a peptide-filled binding cleft and β_2m but without the α3 domain and therefore probably without the ability to bind CD8.

1. Germain, R. N. & Margulies, D. H. (1993) *Annu. Rev. Immunol.* **11**, 403–450.
2. Salter, R. D., Norment, A. M., Chen, B. P., Clayberger, C., Krensky, A. M., Littman, D. R. & Parham, P. (1989) *Nature (London)* **338**, 345–347.
3. Rosenstein, Y., Ratnofsky, S., Burakoff, S. J. & Herrmann, S. H. (1989) *J. Exp. Med.* **169**, 149–60.
4. Blue, M. L., Craig, K. A., Anderson, P., Branton, K., Jr., & Schlossman, S. F. (1988) *Cell* **54**, 413–21.
5. Cerundolo, V., Tse, A. G. D., Salter, R. D., Parham, P. & Townsend, A. (1991) *Proc. R. Soc. London Ser. B* **244**, 169–177.
6. Giblin, P. A., Leahy, D. J., Mennone, J. & Kavathas, P. B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1716–20.
7. Olsen, A. C., Pedersen, L. Ø., Hansen, A. S., Nissen, M. H., Olsen, M., Hansen, P. R., Holm, A. & Buus, S. (1994) *Eur. J. Immunol.* **24**, 385–392.
8. Townsend, A., Elliot, T., Cerundolo, V., Foster, L., Barber, B. & Tse, A. (1990) *Cell* **62**, 285–295.
9. Lehman-Grube, F., Dralle, H., Utermohlen, O. & Lohler, J. (1994) *J. Immunol.* **94**, 595–603.
10. Elliott, T., Elvin, J., Cerundolo, V., Allen, H. & Townsend, A. (1992) *Eur. J. Immunol.* **22**, 2085–2091.
11. Garboczi, D. N., Hung, D. T. & Wiley, D. C. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3429–3433.
12. Guo, H.-C., Jardetzky, T. S., Garrett, T. P. J., Lane, W. S., Strominger, J. L. & Wiley, D. C. (1992) *Nature (London)* **360**, 364–366.
13. MacFerrin, K. D., Chen, L., Schreiber, S. L. & Verdine, G. L. (1993) *Methods Enzymol.* **217**, 79–102.
14. Parker, K. C. & Wiley, D. C. (1989) *Gene* **83**, 117–124.
15. Silver, M., Parker, K. & Wiley, D. C. (1991) *Nature (London)* **350**, 619–622.
16. Silver, M. L., Guo, H.-C., Strominger, J. L. & Wiley, D. C. (1992) *Nature (London)* **360**, 367–369.
17. Watowich, S. J., Skehel, J. J. & Wiley, D. C. (1994) *Acta Crystallogr. D*, in press.
18. Daresbury Laboratory (1979) CCP4 (Collaborative Computing Project No. 4), A Suite of Programs for Protein Crystallography (Sci. Eng. Res. Council, Warrington, U.K.).
19. Navaza, J. (1994) *Acta Crystallogr. A* **50**, 157–163.
20. Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr.* **21**, 273–278.
21. Brünger, A. (1992) X-PLOR Manual (Yale Univ., New Haven, CT), Version 3.0.
22. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47**, 110–119.
23. Brünger, A. T. (1992) *Nature (London)* **355**, 472–475.
24. Monzingo, A. F., Collins, E. J., Ernst, S. R., Irvin, J. D. & Robertus, J. D. (1993) *J. Mol. Biol.* **233**, 705–715.
25. Colovos, C. & Yeates, T. O. (1993) *Protein Sci.* **2**, 1511–1519.
26. Luthy, R., Bowie, J. U. & Eisenberg, D. (1992) *Nature (London)* **356**, 83–85.
27. Hodel, A., Kim, S.-H & Brunger, A. T. (1992) *Acta Crystallogr. A* **48**, 851–858.
28. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1991) *Nature (London)* **353**, 321–325.
29. Stern, L. J. & Wiley, D. C. (1994) *Structure* **2**, 245–251.
30. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257**, 919–927.
31. Madden, D. R., Garboczi, D. N. & Wiley, D. C. (1993) *Cell* **75**, 693–708.
32. Young, A. C. M., Zhang, W., Sacchettini, J. C. & Nathenson, S. G. (1994) *Cell* **76**, 39–50.
33. Ruppert, J., Sidney, J., Celis, E., Kubo, R. T., Grey, H. M. & Sette, A. (1993) *Cell* **74**, 929–937.
34. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. (1991) *Nature (London)* **253**, 326–329.
35. Guo, H.-C., Madden, D. R., Silver, M. L., Jardetzky, T. S., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8053–8057.
36. Ezquerra, A., Bragado, R., Vega, M. A., Strominger, J. L., Woody, J. & Lopez de Castro, J. A. (1985) *Biochemistry* **24**, 1733–1741.
37. Saper, M. A., Bjorkman, P. J. & Wiley, D. C. (1991) *J. Mol. Biol.* **219**, 277–319.
38. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.