The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3

Partho Ghosh*, Miguel Amaya†, Elizabeth Mellins† & Don C. Wiley*‡§

* Department of Molecular and Cellular Biology, and ‡ Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA
† Children's Hospital of Philadelphia and Children's Seashore House, Philadelphia, Pennsylvania 19104, USA

A complex between HLA-DR3 and a fragment of invariant chain called CLIP was isolated from a human cell line defective in antigen presentation and its X-ray crystal structure determined. Previous data indicate that this complex is an intermediate in class II histocompatibility maturation, occurring between invariant chain–DR3 and antigenic peptide–DR3 complexes. The structure shows that the CLIP fragment binds to DR3 in a way almost identical to that in which antigenic peptides bind class II histocompatibility glycoproteins. The structure is the substrate for the loading of antigenic peptides by an exchange process catalysed by DM.

CLASS II histocompatibility membrane glycoproteins (αβ) bind antigenic peptides in endosomal compartments of antigen-presenting cells for 'presentation' at the cell surface to T cells. A second membrane glycoprotein, invariant chain (Ii), complexes with αβ in the endoplasmic reticulum to stabilize it in the absence of a bound peptide and to target it to the endocytic pathway (reviewed in ref. 1). Ii, which blocks the binding of antigenic peptide to αβ, is removed from αβ by stepwise proteolysis in endosomes before antigenic peptide loading. In several cells defective in antigen presentation, complexes of αβ with a nested set of 20–24 residue Ii fragments (within residues 81–104) called CLIP (class II associated invariant chain peptide) accumulate at the cell surface[2–5], raising the possibility that αβ-CLIP is an intermediate in class II molecule maturation[2,4–6]. The observations that *in vitro* proteolysis of αβIi isolated from cells generates αβ-CLIP, and that pulse-chase experiments in cells demonstrate the transient formation of αβ-CLIP, are further evidence that αβ-CLIP is an intermediate that is generated naturally during class II maturation[6,7].

The CLIP segment of Ii has been implicated in several of the activities of Ii. Genetic truncation and deletion experiments[8,9] demonstrate that the CLIP segment is necessary for the *in vivo* activities that promote αβ assembly and stabilize αβ dimers against permanent aggregation with chaperones in the endoplasmic reticulum. The CLIP segment is also necessary for the *in vitro* Ii activity of inhibiting peptide binding to αβ[9,10]. Differing affinities of CLIP for different alleles of DR and analysis of the binding of substituted CLIPs[6,11–13] provide compelling evidence that CLIP binds in the αβ peptide binding site. However, alternative allosteric models have been proposed[8], and studies of a mutant class II molecule[14] and the inhibition motif of a short (92–104) fragment of CLIP[15] have been interpreted to favour an alternative mode of binding.

Various cell lines deficient in antigen presentation have been shown to lack HLA-DM[16,17], which is an MHC-encoded protein. The accumulation of DR3-CLIP on the surface of these cells suggests that DM is required to remove CLIP from DR3 in endosomes so that antigenic peptides are able to bind[16,17]. DM has recently been shown to stimulate the release of CLIP from class II MHC *in vitro*[18–20]. The peptide release effect of DM is limited to CLIP and to at least one other peptide[18], and does not extend to all peptides.

We have purified DR3-CLIP from a cell line deficient in antigen presentation[21] that lacks DM[16,17], and determined its structure from crystals grown at the pH (4.5) of endosomes. The structure reveals that CLIP occupies the peptide binding site and binds in a way almost identical to the way in which a peptide antigen, influenza virus haemagglutinin (HA 306–318), binds to DR1 (ref. 22). The contacts between DR3 and CLIP suggest that the CLIP segment of Ii would stabilize most class II molecules almost as much as an antigenic peptide. We measured the half-life of CLIP bound to DR3 (at 37 °C, pH 4.5) to be almost 4 hours, long enough to require an exchange process, such as that mediated by DM, to remove CLIP from most subtypes of class II MHC before antigenic peptides can bind *in vivo*. The conformation of DR3 in the DR3-CLIP complex is only slightly different from that of DR1 in DR1-HA, although it is sufficient to explain antigenic differences without recourse to allosteric transitions or novel conformations. We show that the structure observed here is the substrate for the DM exchanger, limiting the possible mechanisms for the catalytic activity of DM.

CLIP occupies the peptide binding site

The X-ray crystal structure of DR3-CLIP was determined by molecular replacement. Details of the structure determination (Fig. 1*a*) and refinement are presented in Table 1.

The peptide binding site of DR3 is filled with a continuous segment of electron density, indicating that CLIP binds in the site normally occupied by peptide antigens[22–25] (Fig. 1*b*). CLIP does not appear to bind to other sites on DR3, as no substantial electron density is observed that is not due to DR3, carbohydrate or water. The excellent quality of the peptide electron density shows that the DR3 is occupied almost exclusively by CLIP. Of the 24 residues in CLIP (L81 M104), 15 are visualized in the binding site (P87 A101), whereas 4–5 residues at the amino terminus and 2–3 residues at the carboxy terminus extend out of the site and are presumably disordered in the crystal (Fig. 1*b*; Table 2).

The composition, stability and ability for peptide exchange of DR3-CLIP, which was derived from DM-deficient cells and solubilized with papain, are comparable to those of DR3-CLIP complexes reported previously[2–5]. Three peptides have been shown by mass spectroscopy to predominate in both DR3 recovered from crystals and in DR3 solubilized by papain digestion before crystallization: P82–M104, P82–P103, and K83–P103 (Fig. 1*b*; Table 2). They are among the major forms of CLIP eluted from DR3 on the cell line deficient in antigen presentation[4,5]. Only small quantities of peptides with the first amino acid of CLIP (81L) were observed after papain

§ To whom correspondence should be addressed.

TABLE 1 Structural analysis and refinement

Crystallographic data			Refinement results	
Resolution (Å)	15–2.75	3–2.75		6–2.75
Independent reflections	13,021	2,876	Reflections	11,772
Multiplicity	3.2	3.0	Atoms	3,095
Completeness (%)	96.9	97.2	R_{free} †	32.5%
Average $I/\sigma(I)$	6.7	2.3	R_{cryst} †	24.6%
R_{merge} (%)	7.1	31.4	r.m.s. deviations	
			Bonds (Å)	0.015
			Angles (deg) ‡	1.57
			Dihedrals (deg) ‡	24.6
			Impropers (deg) ‡	1.31
			B-factors (Å^2) §	4.6
			Real space fit ‖	0.85 ± 0.06
			Average *B*-factors	
			HLA-DR3 (Å^2)	47 ± 18
			CLIP (Å^2)	49 ± 16

Purification and crystallization. HLA-DR3 was purified[4] from 9.5.3 cells[21], solubilized in β-octylglucoside, and cleaved with papain[40]; it was further purified using a BioSil S3000 column and concentrated to ~15 mg ml^{-1}. Crystals were grown using conditions similar to those reported for other DR molecules[41]: 12% PEG 4K, 100 mM $MgCl_2$, 100 mM acetate buffer, pH 4.5. Crystals belong to space group $P3_212$ with cell dimensions $a = b = 78.5$ Å, $c = 159.1$ Å. **Data collection and processing.** Diffraction was collected as oscillations of 1° (CHESS F-1; $\lambda = 0.91$ Å) on Fuji phosphor plates from two flash-frozen crystals (100 K); before freezing, crystals were soaked for 5 min in 15% glycerol, 15% ethylene glycol, 16% PEG 4K, 100 mM $MgCl_2$, 100 mM acetate, pH 4.5; mosaicity of frozen crystals was ~0.2°. Data were indexed and integrated using DENZO (Z. Otwinowski, personal communication), and scaled and reduced using ROTAVATA/AGROVATA, POSTREF and TRUNCATE[42]. Partial reflections were summed. **Molecular replacement.** A molecular replacement solution was identified with Amore[43] using a single αβ DR1 (ref. 22) heterodimer, for which 19 residues differing between DR1 and DR3 were substituted by Ala, as the model. The rotation (15–3 Å) and translation (8–4 Å) solutions were the highest peaks in their respective maps. After rigid-body refinement (8–3.2 Å, with Amore), the correlation coefficient (c.c.) of the rotation and translation solution was 0.59 and the R_{cryst} 42.7% (next-highest solution: c.c., 0.27; R_{cryst} 54.8%). The translation search identified $P3_212$ as the correct enantiomorph. The asymmetric unit of the crystal contains a single αβ heterodimer of DR3, and the same dimer of αβ heterodimers observed previously[22–25] was observed here as a crystallographic dimer. **Model building and refinement.** A random 10% of reflections were omitted from refinement for R_{free}[44] calculation. Before refinement, the model B-factor was set to 45 Å^2 (initial R_{free}, 44.1%; R_{cryst}, 45.7%, 10–3.5 Å). The model was rigid-body refined using X-PLOR[44], and residues specific to DR3 were built into $2F_o - F_c$ omit maps (15–3.5 Å) using O[45]. Preferred rotamer conformations were used when possible, and a database of refined structures was used for main-chain rebuilding[46]. Only procedures which minimized R_{free} were followed; the difference between R_{free} and R_{cryst} was kept as small as possible. After a single rebuilding and refinement step (positional and group *B*-factor, 6–3.5 Å), an anisotropic *B*-factor tensor ($B_{11} = B_{12} = B_{22} = -10$ Å^2, and $B_{33} = 10$ Å^2) determined from a Wilson plot was applied, resulting in a 1% drop in R_{free}. The resolution limit was increased gradually to 2.75 Å through subsequent rebuilding and refinement cycles, which included positional and restrained atomic *B*-factor refinement, geometric regularization, and occasionally simulated annealing. The model was rebuilt into weighted (with SIGMAA[42]) $2F_o - F_c$ omit maps (15–2.75 Å), in which either approximately one-tenth of the model or entire protein domains were omitted, or into simulated annealing omit maps. The peptide was modelled into $F_o - F_c$ density after 6 cycles of rebuilding and refinement (R_{free}, 34.8%; R_{cryst}, 27.8%; 6–2.75 Å). After 10 cycles (R_{free}, 33.4%; R_{cryst}, 23.8%; 6–2.75 Å), waters within 2.6–3.4 Å of a hydrogen bonding donor or acceptor were built into $\geq 3\sigma$ $F_o - F_c$ density. Positionally unstable waters or those with refined *B*-factors greater than 60 Å^2 were removed. The final model contains 22 waters. Only one *N*-acetylglucosamine monosaccharide unit could be fit at the N-linked glycosylation sites, α78N and α118N. No carbohydrate could be modelled at the third N-linked glycosylation site, β19N. All residues have allowed φ, ψ angles, and coordinate error is estimated to be 0.35–0.45 Å[46]. $2F_o - F_c$ maps show clear electron density for residues 5–180 of the α-subunit and residues 5–191 of the β-subunit. Three small breaks in the main-chain density occur in loop regions of the β_2 domain (at β109, β172 and β189). The side chains of 15 residues on surface loops are disordered (α123R, β19N, β22E, β23R, β105K, β107Q, β109L, β110Q, β111H, β136Q, β138E, β166R, β167S, β189R and β191R). These residues are modelled, but atoms beyond Cβ were omitted in the later refinement stages. A break occurs in the side-chain density of β74R between Cδ and Cζ. Although the guanidinium of β74R is positioned in electron density, this density may instead correspond to a solvent molecule. The side-chain amide and carboxyl groups of Q100 of CLIP are not found in electron density, and therefore calculations of buried surface area do not include these two groups.

* $R_{merge} = (\sum_h \sum_i |I_{h,i} - I_h|)/\sum_h \sum_i I_{h,i}$, where I_h is the mean intensity of symmetry-related reflections, $I_{h,i}$.

† $R_{value} = (\sum_{hkl} |F_{obs} - F_{calc}|)/\sum_{hkl} F_{obs}$, where R_{free} is calculated for a randomly chosen 10% of reflections ($F > 0$) omitted from refinement, and R_{cryst} is calculated for the remaining 90% of reflections ($F > 0$) included in refinement.

‡ Root-mean-square deviations from ideal values.

§ Root-mean-square deviations between *B*-factors of bonded atoms.

‖ Real-space fit correlation coefficient[45] calculated from $2F_o - F_c$ electron density and the refined model of DR3-CLIP.

TABLE 2 Nature and stability of DR3-CLIP isolated from DM-deficient cells

		Observed *m/z*	Calculated *m/z*
Peptides isolated from crystals of DR3	P82–M104	2563.41	2564.27
	P82–P103	2432.15	2433.08
	K83–P103	2334.82	2335.96
		Conditions	
Half-life of DR3-CLIP †	3.7 h	pH 4.5, 37 °C	
Effect of HLA-DM ‡	Exchange of CLIP	pH 4.5, 37 °C	
SDS-stability of DR3-CLIP §	Unstable	1% SDS, 25 °C	

* Masses of peptides, which were acid-extracted[35] from papain-cleaved DR3 and from washed and solubilized crystals of DR3, were determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry[36]. The three peptides above were the predominant ones in mass analyses of pre- and postcrystallization samples. Because of PEG suppression in the postcrystallization samples, peptides that were of very low abundance in the precrystallization sample were not detectable in the postcrystallization sample. The shortest peptide detected was K83-L102, and a peptide beginning at the first residue of CLIP, L81-P103, was also detected, both in very low quantities and only in the precrystallization sample.

† Soluble DR3 at a concentration of 2 μM was incubated in 100 mM acetate buffer, pH 4.5 at 37 °C for various times and monitored by absorbance, A_{280}, for quantity of DR3-CLIP complex remaining using high-performance liquid chromatography (HPLC) gel filtration (BioSil S3000); tyrosine was included as an internal standard. The initially homogeneous DR3-CLIP gel filtration peak yielded over time a separable mixture of aggregates and DR3-CLIP complexes; the aggregates are probably an indication of peptide dissociation and the formation of 'empty' molecules[29]. The half-life was determined from a first-order exponential fit of the data between 0 and 300 min; beyond 300 min, the amount of DR3-CLIP remains at a steady-state level. Detergents such as β-octylglucoside substantially decreased the DR3-CLIP half-life as previously reported[6], as did Zwittergent (data not shown), which was used in studies reporting short half-lives of DR2-CLIP complexes[15].

‡ Papain-solubilized DR3, from DM-deficient cells, at a concentration of 20 nM was incubated with or without 4 μM soluble, recombinant HLA-DM[18] and with 5 μM biotin-labelled IgCk (37–51) (KVQWKVDNALQSGNS) peptide for 4 h at pH 4.5, 37 °C (D. Zaller, personal communication). Incubation with DM led to a threefold increase in the amount of CLIP exchanged for IgCk peptide, similar to results reported for full-length DR3-CLIP (compare with Fig. 4*a* in ref. 18).

§ DR3-CLIP from samples used for crystallization and from solubilized crystals were unstable in SDS–PAGE in the absence of boiling, yielding monomers rather than heterodimers.

treatment (Table 2), but this should have no structural consequences because CLIP with the first two residues (81–82) missing has the same stability on DR[26] and participates in the DM exchange reaction in the same way as CLIP (81–104)[19]. Furthermore, the soluble DR3-CLIP complexes used for crystallization were shown to participate in the DM exchange reaction (D. Zaller, personal communication), similar to complexes purified without papain treatment from DM-deficient cells (Table 2).

We also observed that DR3-CLIP complexes after papain digestion and resolubilized from crystals have the characteristic instability to sodium dodecylsulphate (SDS) (Table 2) of DR3-CLIP reported previously[2,3].

Interactions between CLIP and DR3

The binding of CLIP to DR3 is remarkably similar to the binding of both the influenza virus HA peptide[22] and a collection of self-peptides[23,24] to DR1. Bound CLIP is extended in a polyproline type II conformation[22] (Fig. 2*a*); it forms a nearly identical, extensive network of hydrogen bonds to conserved class II MHC residues[22] (Fig. 2*a*); and its side chains extend into the

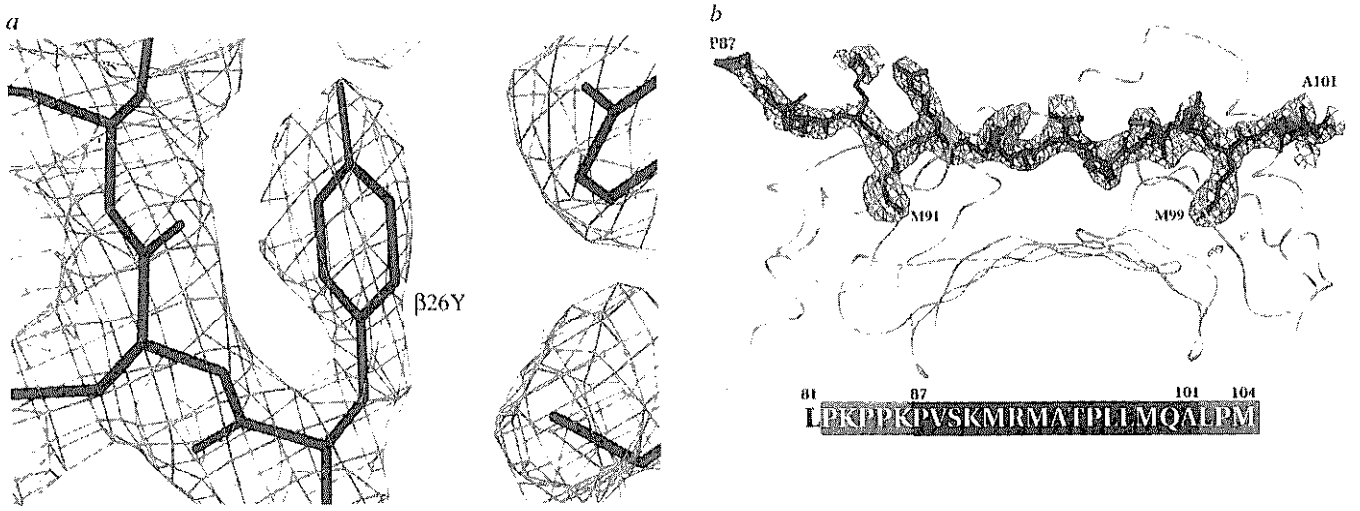

FIG. 1 Electron density for CLIP. *a*, Electron density for β26Y, a residue differing between DR3 and DR1 (β26L), is clearly present in a $2F_o - F_c$ map (15–2.75 Å, contoured at 1σ) calculated before any rebuilding or refinement (except for rigid-body) of the molecular replacement model. In the model, β26 is replaced by Ala (see Table 1 legend). This figure and all subsequent figures, except for 2*b*, were made with MidasPlus[47]. *b*, Difference electron density shows that 15 residues of CLIP (P87-A101) are visible in the peptide binding site. Two methionines, M91 and M99, form the prominent density near the ends of CLIP and point down towards the floor of the binding site. The electron density is from an omit $F_o - F_c$ map (15–2.75 Å, contoured at 2σ) calculated from the refined model of DR3-CLIP from which CLIP was omitted. The $\alpha_1\beta_1$ peptide-binding domain (α5–α80, β5–β94) of DR3 is shown in a wire trace with the β-sheet floor of the binding site at the bottom, the α_1 domain α-helix in back, and the β_1 domain α-helix in front. Below the wire trace is the sequence of CLIP (L81-M104) with residues observed in the crystal structure shown in the central black box and disordered residues not observed in the structure shown in the surrounding grey boxes.

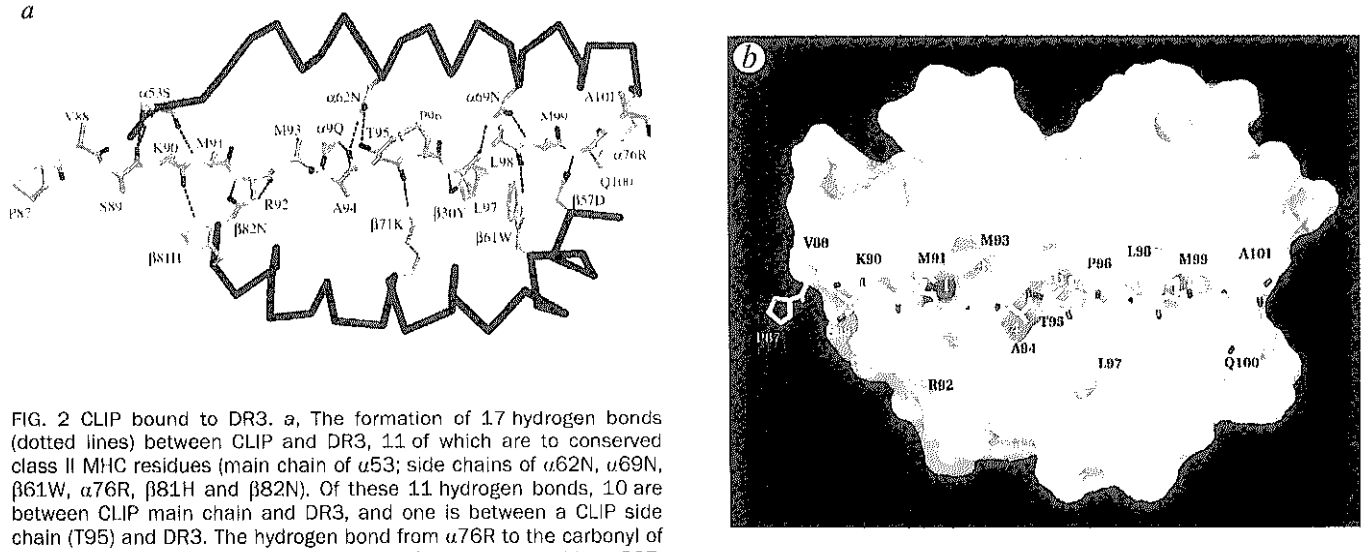

FIG. 2 CLIP bound to DR3. *a*, The formation of 17 hydrogen bonds (dotted lines) between CLIP and DR3, 11 of which are to conserved class II MHC residues (main chain of α53; side chains of α62N, α69N, β61W, α76R, β81H and β82N). Of these 11 hydrogen bonds, 10 are between CLIP main chain and DR3, and one is between a CLIP side chain (T95) and DR3. The hydrogen bond from α76R to the carbonyl of the CLIP residue Q100 is partly obscured. CLIP (yellow) residues P87, T95 and P96 are shown fully, whereas other CLIP residues are displayed only up to Cβ. Only DR3 (silver) residues making hydrogen bonds with CLIP are shown. The view is from above the binding site with the α_1 domain α-helix at the top and the β_1 domain α-helix (both in red Cα trace) at the bottom. *b*, Six side chains of CLIP fit in pockets formed by DR3. CLIP is shown as bonds whereas DR3 is shown as a surface. View as in *a*. Contacts with DR3 bury 58% of the surface area of CLIP P87-A101. Pockets 1 and 9 bury 190 and 175 Å^2 of CLIP solvent-accessible surface area, respectively; pockets 6 and 7, smaller cavities, bury 122 and 160 Å^2; pocket 3, buries only 118 Å^2, and pocket 4 buries 99 Å^2. The methylene group of arginine 92 of CLIP, which does not fit into a pocket, nevertheless contacts the surface of DR3 and positions the guanidinium group 3.2 Å from β81 His.

METHODS. Calculation of solvent-accessible surface areas were performed with a probe radius of 1.4 Å. Residues contacting CLIP in each pocket are: pocket 1, α24F, α31I, α32F, α43W, α52A, main chain of α53S, α54F, β81H, β82N, β85V, β86V; pocket 3, α9Q, α22F, α24F, α54F, α55E, α58G, α59A, α61A, α62N, β78Y; pocket 4, α9Q, α62N, β13S, β74R, β78Y; pocket 6, α9Q, α11E, α62N, α65V, α66D, α69N, β11S, β13S, β28D, β30Y, β71K, β74R; pocket 7, α65V, α69N, β30Y, β47F, β61W, β67L, β70Q, β71K, β74R; pocket 9, α69N, α72I, α73M, α76R, β9E, β30Y, β37N, β38V, β57D, β61W. R92 of CLIP does not form a pocket contact, but 121 Å^2 of its surface area is rendered inaccessible to solvent by contact with DR3. This figure was made with GRASP[48].

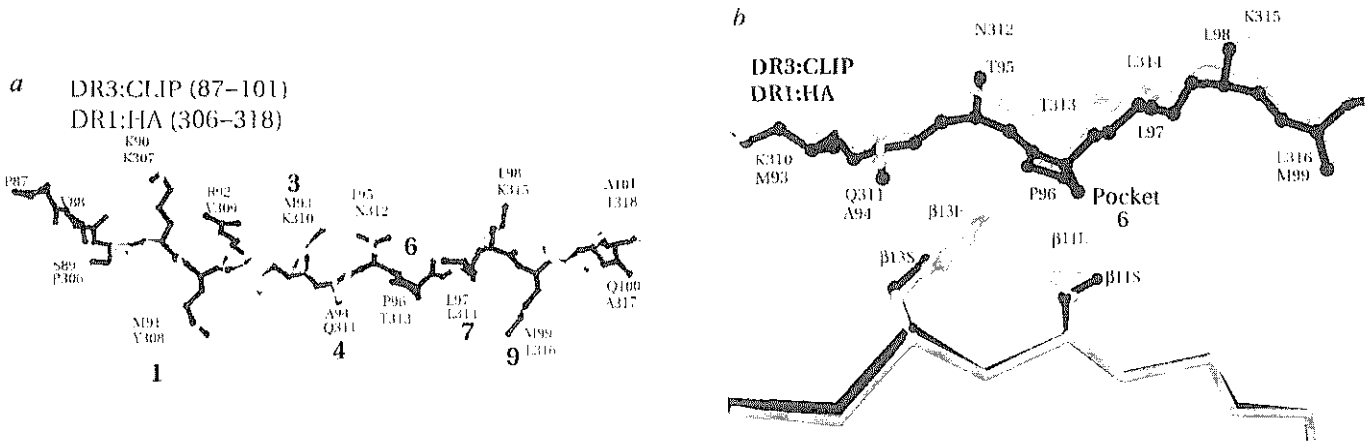

FIG. 3 Conformation of CLIP. *a*, Superposition of CLIP 87-101 with HA 306-318, based on least-squares α-carbon overlap[49] of the $\alpha_1\beta_1$ peptide-binding domains of DR3 and DR1. Both peptides have polyproline type II conformations, with mean ϕ, ψ angles of $-87\pm20°$, $140\pm14°$ for CLIP, and $-82\pm15°$, $139\pm16°$ for HA. View as in Fig. 1*b*. The positions of pockets are indicated. CLIP extends two residues further at the N terminus than HA. Although the first 5 overlapping residues (S89-M93 of CLIP, P306-K310 of HA) have similar main-chain paths (r.m.s. deviation of 0.31 Å), the last 8 residues (A94-A101 of CLIP, Q311-T318 of HA) differ (r.m.s. deviation of 1.12 Å), with the largest displacement at pocket 6. *b*, Superposition of DR3-CLIP (red) and DR1-HA (blue) at pocket 6 showing the displacement of P96 of CLIP and the allelic differences between DR3 and DR1 at β11 and β13. The peptides are represented as main-chain traces with the position of side chains indicated by their Cβ atoms; main-chain carbonyls are not shown. Side chains for the residues in pocket 6, P96 for CLIP and T313 for HA peptide, are shown. The side chains for β11 and β13 are shown in a ball-and-stick model, and the β-strand containing these residues is shown in a Cα trace.

same pockets in the DR3 binding site as those observed in DR1 (ref. 22) (Fig. 2*b*).

Hydrogen bonds

The conformation of the main chain of bound CLIP is extended but twists such that side chains project from the main chain about every 128°, as in DR1-HA. This allows 17 hydrogen bonds to form between atoms of CLIP and DR3 (Fig. 2*a*), of which 11 are to MHC atoms conserved in class II MHC sequences. The ten hydrogen bonds between the CLIP main-chain atoms and DR3 were observed in the DR1-HA complex and predicted to play a universal role in binding peptides to class II molecules[22]. Overall, the pattern of 17 hydrogen bonds in DR3-CLIP is essentially identical to the 18 bonds in the DR1-HA complex[22] (compare Fig. 2*a* here and Fig. 3 from ref. 22).

CLIP side chains and DR3 pockets

A series of pockets that accommodate the side chains of CLIP in the DR3 peptide binding site (labelled 1, 3, 4, 6, 7 and 9 in Fig. 2*b*) are positioned in the same way as those in DR1 (ref. 22). Pockets 1 and 9, deep cavities at either ends of the site, hold CLIP methionines 91 and 99 (Fig. 2*b*); pockets 6 and 7, smaller cavities in the centre of the site, hold CLIP proline 96 and leucine 97; and pocket 3, which is more like a non-polar shelf, holds CLIP methionine 93 and leaves 30% of the methionine's surface exposed. In the DR1-HA complex, that pocket holds the methyl-

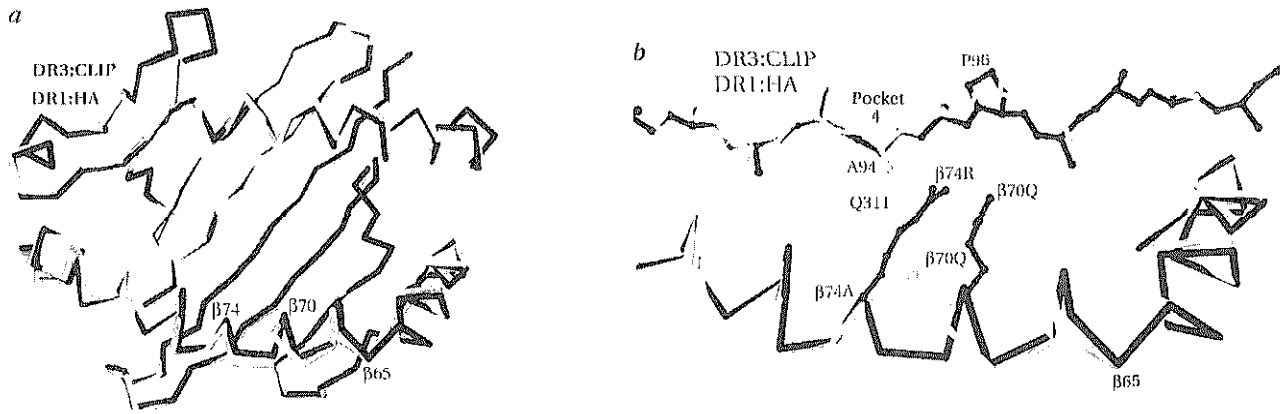

FIG. 4 A comparison of DR3 and DR1 from the CLIP and HA complexes. *a*, Superposition of the $\alpha_1\beta_1$ peptide-binding domains of DR3-CLIP (red) and DR1-HA (blue) (residues α5–α80 and β5–β94). View as in Fig. 2. The $\alpha_1\beta_1$ peptide binding domains are very similar except for an α-helical region between β65 and β74. (Not shown are the immunoglobulin-like α_2 and β_2 domains which differ between the two complexes by small rotations and translations, probably as a result of different crystal packing forces. Relative to DR1-HA, the α_2 domain of DR3-CLIP is rotated 3° and translated 1 Å, and the β_2 domain is rotated 5° and translated 2 Å, as determined by a least-squares superposition based on the Cα positions of the $\alpha_1\beta_1$ peptide-binding domains. Individually, the α_2 domains of DR3 and DR1 superimpose with an r.m.s. deviation of 0.56 Å and the β_2 domains with an r.m.s. deviation of 0.63 Å. Similar shifts in the immunoglobulin-like domains have been observed for class I MHC[50]). *b*, Superposition of DR3-CLIP (red) with DR1-HA (blue) at the β65–β74 region of the β_1 domain α-helix. View as in Fig. 2. The main-chain atoms of the peptides are shown, except for main-chain carbonyls. Side-chain positions are indicated by the Cβ atoms. The side chains of Q311 of HA peptide and for reference, P96 of CLIP, are shown.

TABLE 3 MHC class II pockets involved in binding CLIP

Conserved pockets	CLIP residue	Area conserved (%) *	DR3 residues: conserved or conservatively substituted in class II MHC †
1	M91	75	**α32F, α43W,** main chain α53S, **α54F, β81H, β82N,** β85 (V/A/L)
3	M93	60	α22 (Y/F), **α54F,** α55 (D/E), **α62N,** β78 (Y/F/M/I)
7	L97	58	α65 (V/I), **α69N,** β47 (Y/F), **β61W,** β67 (L/I/F/V)
9	M99	54	**α69N,** α73 (M/L/A/V), **α76R,** β38 (V/L/A), **β61W**
Variable pockets			
6	P96	41	**α62N,** α65 (V/I), **α69N**
4	A94	23	**α62N,** β78 (Y/V/M/I)

* Percentage of DR3-accessible surface area buried by contact with side-chain atoms of CLIP that is composed of residues conserved or conservatively substituted in class II MHC sequences. A probe radius of 1.4 Å was used to calculate solvent-accessible surface area (M. Handschumacher and F. Richards, personal communication). The surface area of DR3 buried by each residue of CLIP was calculated and partitioned with respect to whether the buried DR3 residue was conserved or conservatively substituted as opposed to non-conserved in class II MHC sequences. For example, of the 87.6 Å^2 of DR3 buried by side-chain atoms of M91 in pocket 1, 65.4 Å^2 or 75% derives from residues conserved or conservatively substituted in class II MHC sequences, while the remaining 22.2 Å^2 derives from residues non-conserved in class II MHC sequences.

† Residues conserved (in bold) or conservatively substituted in class II MHC sequences that are buried in the DR3-CLIP complex by contact with side-chain atoms of CLIP. After each conservatively substituted residue is indicated the range of substitutions found in human class II MHC.

ene groups of a peptide lysine[22]. Pocket 4 of DR3 holds the side-chain methyl group of CLIP Ala 94. This pocket appears to have collapsed slightly (see below) relative to the size it may exhibit when holding an aspartic acid side chain in the common DR3 peptide motif[27,28]. The register of CLIP with respect to the protein pockets is in accordance with that proposed from substitution studies of CLIP[11,13].

Peptides that bind to DR3 exhibit sequence preferences at position 1 where non-polar residues are preferred (M, L, V, I, F, Y), and position 4 where negatively charged aspartic acid is preferred[27,28]. This binding motif correlates well with the non-polar nature of pocket 1 and the positively charged arginine (β74) in pocket 4 of DR3. Small rearrangements must be possible in both pockets as model building indicates that tyrosine (Y) could not fit in pocket 1 or aspartic acid (D) in pocket 4 without slight movements of DR3 side chains. Some peptides that bind DR3 exhibit a second motif such that, when the residues in pockets 1 and 4 are not ideal (A in 1; N, Q, E, S, T in 4), peptide position 6 is often positively charged (K, H, R). This appears to correlate with the negative charge in pocket 6 from β28 aspartic acid of DR3.

Comparison of DR3-CLIP with DR1-HA

Overall, CLIP appears to fit as closely against DR3 as the HA peptide does against DR1. Buried in the CLIP-DR3 complex are 1,386 Å^2 of solvent-accessible surface area of CLIP and 904 Å^2 of DR3, comparable to the 1400 Å^2 of the HA peptide and 930 Å^2 of the DR1 in DR1-HA[22].

The path of the peptide main chain is nearly identical for CLIP and HA peptide in the first half of the MHC molecule's binding site, but diverges slightly in the second half at pocket 6 (Fig. 3*a*), where the main chain of CLIP is 1.8 Å closer to the β-sheet floor of the binding site. The difference of two small amino acids, serines at β11 and β13 in DR3, for two large residues, leucine (β11) and phenylalanine (β13) in DR1, allows the peptide in DR3 to get slightly deeper into the site (Fig. 3*b*). A proline, CLIP P96, in pocket six also contributes as its cyclic side chain does not protrude from the peptide backbone.

There is only one small conformation difference in the DR backbone structures between DR3-CLIP and DR1-HA. Overall the $\alpha_1\beta_1$ peptide-binding domains superimpose very well with an r.m.s. deviation of 0.51 Å in α-carbon positions (Fig. 4*a*). The small difference occurs in the β-chain α-helix between residues β65 and β74. The result of differences in the locations of β70 and β74 in DR1 and DR3 is that the β-chain α-helix appears to collapse slightly (1.5 Å) towards the peptide in DR3-CLIP, probably because the peptide side chain in pocket 4 is small (Fig. 4*b*). Modelling an aspartic acid at peptide position 4 requires that β74R is moved to open pocket 4 slightly.

Complexes with other class II MHCs

Many of the same binding interactions observed in the DR3-CLIP complex are probably found in CLIP complexes with other class II MHC molecules. The four large CLIP side chains M91, M93, L97 and M99 are bound in those DR3 pockets which have the greatest proportion of their buried surface areas formed by residues conserved or conservatively substituted in class II molecules (Table 3). Thus these CLIP residues may fit into most class II binding sites. The two smallest CLIP side chains, A94 and P96, are bound in those pockets with linings that are the most variable in chemical character in class II molecules. Pocket 4, which holds A94 of CLIP, is extremely variable (23% conserved). Similarly, 41% of the surface area of pocket 6, which holds P96, is formed by residues that undergo non-conservative substitutions in class II molecules. The occurrence of small side chains at positions 94 and 96 in CLIP (A94, P96) may minimize unfavourable contacts in pockets 4 and 6 of most other class II molecules.

Implications for Ii activities

After proteolytic degradation of Ii *in vivo*, CLIP is seen, from the structure of DR3-CLIP, to bind in the αβ peptide binding site using the same network of hydrogen bonds and contacts to protein pockets as an antigenic peptide. This provides a mechanism for the chaperoning role of Ii, because CLIP in the binding site would be expected to stabilize αβ against aggregation *in vivo*[9], as antigenic peptides do *in vitro*[29,30]. Genetic truncation experiments[8,10], pulse-chase experiments[6], antibody binding[31] (E.M. and Y. Paterson, manuscript in preparation) and T-cell[31] recognition experiments have already provided evidence that Ii binds to MHC class II molecules by inserting CLIP into the peptide binding site. The region surrounding CLIP in a soluble, recombinant Ii has recently been shown to be unstructured, as determined by nuclear magnetic resonance and proteolytic sensitivity[32,33], making CLIP available for binding as observed here. The presence of CLIP in the peptide binding site also explains how Ii blocks peptide binding *in vitro*[9,10]. For those class II molecules (for example, I-A^k and DR52a) which have a low affinity for CLIP[12], the C-terminal domain of Ii may augment the affinity[33].

Most studies of CLIP binding to DR[11,13,34] have suggested the binding mode observed here, and the effects on CLIP binding by mutations at α58 and β86 in a mouse class II MHC[14] are also consistent with the observation that CLIP binds near both these residues (Fig. 2*b* legend). A suggestion that the C-terminal part of CLIP makes contacts to DR2 other than those made by conventional groove-binding peptides or as described for CLIP[11,13,34] could have resulted from the use of a short peptide lacking the P1 anchor and/or the addition of an N-terminal fluorescent probe to that peptide[15].

The details of the atomic contacts between CLIP and DR3 are consistent with the *in vivo* stability inferred from the ability of DR3-CLIP to traverse the endosome and survive on the cell

surface in DM-deficient and wild-type cells[2,6,35,36]. The half-life of soluble DR3-CLIP isolated from DM-deficient cells is almost 4 hours in conditions (pH 4.5 and 37 °C) mimicking the endosomal environment (Table 2). This is longer than the 1–2 hours required for class II MHC maturation[37], and suggests that the structure of DR3-CLIP represents the structure of the complex in the endosome, before DM-mediated exchange.

CLIP's nine N-terminal residues (L81–S89) have been suggested to increase CLIP dissociation from soluble DR1 (ref. 26) and from detergent-solubilized DR2 (ref. 15). In the latter case, the fast off-rate reported was probably the result of using a detergent (Zwittergent), which has a β-OG-like destabilizing effect[6] on DR3-CLIP (data not shown). In the former case, experiments are needed to determine whether it is the ordered or disordered residues at the terminus that confer this reported property.

Implications for DM exchange reaction

The finding that CLIP binds to DR3 like HA binds DR1 has implications for DM's exchange mechanism[18,20]. Although previous observations of a difference in SDS stability and antibody binding[2,5,21] had been interpreted to indicate an altered DR3-CLIP conformation, the structure shows only a small difference between DR3-CLIP and DR1-HA: a segment of the β-chain α-helix (β65–74) in DR3-CLIP is displaced about 1.5 Å towards the peptide binding site (Fig. 4), forming a slightly more 'closed' site rather than the more 'open' site that SDS instability was thought to indicate. This 'closing' may result from movement of β74 Arg and the collapse of pocket 4 to fit around alanine 94 of CLIP, a peptide position normally occupied by aspartic acid in peptides that bind DR3. Replacement of alanine 94 by aspartic acid results in SDS-stable complexes[11]. The functional consequences, if any, of the lack of SDS stability (and of instability in β-OG[6]) of DR3-CLIP are not clear. Some antigenic peptides, even ones that are recognized by T cells, form complexes that are SDS unstable[30,38] and, importantly, some DR-CLIP complexes, such as DR1-CLIP, are stable in SDS[35] but readily exchanged by DM[18,20]. This segment of DR3's β chain probably also accounts for the difference in antibody recognition between DR3-CLIP and other DR3-peptide complexes, as the binding site of an antibody, NDS-9, that recognizes DR3-peptide complexes but not DR3-CLIP[4] has been mapped to β73, 74 and 77 (ref. 39) (Fig. 4).

The observation that DR3-CLIP and DR1-HA are highly similar in structure suggests that DM does not discriminate between CLIP and HA by binding to an alternate equilibrium conformer of DR that would be present in DR-CLIP but not in DR-HA. Instead, the increase in the rate of peptide dissociation by DM suggests that DM may recognize and stabilize a strained, transition-state-like conformer of DR, lowering the free-energy barrier for CLIP dissociation. This conformer may exist for both DR-HA and DR-CLIP, but stabilization by DM may result in the preferential dissociation of peptides lacking optimal anchor residues, such as CLIP[18,20]. The structure of DR3-CLIP, a substrate for DM-mediated exchange, suggests that DM functions on a transient state of class II MHC rather than on a new equilibrium conformation. □

Received 15 May; accepted 24 October 1995.

1. Cresswell, P. *A. Rev. Immun.* **12,** 259–293 (1994).
2. Riberdy, J. M., Newcomb, J. R., Surman, M. J., Barbosa, J. A. & Cresswell, P. *Nature* **360,** 474–477 (1992).
3. Sette, A. *et al. Science* **258,** 1801–1804 (1992).
4. Mellins, E. *et al. J. exp. Med.* **179,** 541–549 (1994).
5. Monji, T., McCormack, A. L., Yates, J. R. & Pious, D. *J. exp. Med.* **153,** 4468–4477 (1994).
6. Avva, R. & Cresswell, P. *Immunity* **1,** 763–774 (1994).
7. Xu, M., Capraro, G. A., Daibata, M., Reyes, V. E. & Humphreys, R. E. *Molec. Immun.* **31,** 723–731 (1994).
8. Freisewinkel, I. M., Schenck, K. & Koch, N. *Proc. natn. Acad. Sci. U.S.A.* **90,** 9703–9706 (1993).
9. Romagnoli, P. & Germain, R. N. *J. exp. Med.* **180,** 1107–1113 (1994).
10. Bijlmakers, M. E., Benaroch, P. & Ploegh, H. L. *J. exp. Med.* **180,** 623–629 (1994).
11. Malcherek, G., Gnau, V., Jung, G., Rammensee, H.-G. & Melms, A. *J. exp. Med.* **181,** 527–536 (1995).
12. Sette, A., Southwood, S., Miller, J. & Apella, E. *J. exp. Med.* **181,** 677–683 (1995).
13. Geluk, A., van Meijgaarden, K. E., Drijfhout, J. W. & Ottenhof, T. H. M. *Molec. Immun.* **32,** 975–981 (1995).
14. Gautam, A. N., Pearson, C., Quinn, V., McDevitt, H. O. & Milburn, P. J. *Proc. natn. Acad. Sci. U.S.A.* **92,** 335–339 (1995).
15. Kropshofer, H., Vogt, A. B. & Hämmerling, G. J. *Proc. natn. Acad. Sci. U.S.A.* **92,** 8313–8317 (1995).
16. Morris, P. *et al. Nature* **368,** 551–554 (1994).
17. Fling, S. P., Arp, B. & Pious, D. *Nature* **368,** 554–558 (1994).
18. Sloan, V. S. *et al. Nature* **375,** 802–806 (1995).
19. Denzin, L. K. & Cresswell, P. *Cell* **82,** 155–165 (1995).
20. Sherman, M. A., Weber, D. A. & Jensen, P. E. *Immunity* **3,** 197–205 (1995).
21. Mellins, E. *et al. Nature* **343,** 71–74 (1990).
22. Stern, L. J. *et al. Nature* **368,** 215–221 (1994).
23. Brown, J. H. *et al. Nature* **368,** 33–39 (1993).
24. Jardetzky, T. S. *et al. Nature* **368,** 711–718 (1994).
25. Kim, J., Urban, R. G., Strominger, J. L. & Wiley, D. C. *Science* **266,** 1870–1874 (1994).
26. Urban, R. G., Chicz, R. M. & Strominger, J. L. *J. exp. Med.* **180,** 751–755 (1994).
27. Malcherek, G. *et al. Int. Immun.* **5,** 1229–1237 (1993).
28. Geluk, A. *et al. J. immun.* **152,** 5742–5748 (1994).
29. Stern, L. J. & Wiley, D. C. *Cell* **68,** 465–477 (1992).
30. Germain, R. N. & Rinker, A. G. *Nature* **363,** 725–728 (1993).
31. Morkowski, S. *et al. J. exp. Med.* **182,** 1403–1413 (1995).
32. Jasanoff, A., Park, S.-J. & Wiley, D. C. *Proc. natn. Acad. Sci. U.S.A.* **92,** 9900–9904 (1995).
33. Park, S.-J., Sadegh-Nasseri, S. & Wiley, D. C. *Proc. natn. Acad. Sci. U.S.A.* **92,** 11289–11293 (1995).
34. Lee, C. & McConnell, H. *Proc. natn. Acad. Sci. U.S.A.* **92,** 8269–8273 (1995).
35. Chicz, R. M. *et al. Nature* **358,** 764–768 (1992).
36. Chicz, R. M. *et al. J. exp. Med.* **178,** 27–47 (1993).
37. Tulp, A., Verwoerd, D., Dobberstein, B., Ploegh, H. L. & Pieters, J. *Nature* **369,** 120–126 (1994).
38. Nelson, C. A., Petzold, S. J. & Unanue, E. R. *Nature* **371,** 250–252 (1994).
39. Maurer, D. & Gorski, J. *J. Immun.* **146,** 621–626 (1991).
40. Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R. & Strominger, J. L. *J. biol. Chem.* **262,** 16087–16094 (1987).
41. Gorga, J. C., Brown, J. H., Jardetzky, T., Wiley, D. C. & Strominger, J. L. *Res. Immun.* **142,** 401–407 (1991).
42. Collaborative Computational Project, N. *Acta crystallogr.* **D50,** 760–763 (1994).
43. Navaza, J. *Acta crystallogr.* **A50,** 157–163 (1994).
44. Brünger, A. T. *Nature* **355,** 472–475 (1992).
45. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47,** 110–119 (1991).
46. Luzzati, V. *Acta crystallogr.* **5,** 802–810 (1952).
47. Huang, C. C., Petterson, E. F., Klein, T. E., Ferrin, T. E. & Langridge, R. *J. molec. Graphics* **9,** 230–236 (1991).
48. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins* **11,** 281–296 (1991).
49. Kabsch, W. *Acta crystallogr.* **A34,** 827–828 (1978).
50. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Cell* **70,** 1035–1048 (1992).

ACKNOWLEDGEMENTS. We thank P. Rosenthal, E. Collins, L. Stern, G. Weiss, A. Taurog, D. Garboczi, S. Ray, S.-J. Park, J. Kim and the staff of the Cornell High Energy Synchrotron Source MacCHESS for help in data collection; R. Chicz for mass spectrometry; P. Pietras for large-scale tissue-culture growth; E. Petterson for help with MidasPlus; and C. Mueller, F. Rey, A. Jasanoff, S.-J. Park and others for discussion. P.G. was supported by the Irvington Institute for Medical Research; M.A. and E.M. by NIH; and D.C.W. is an investigator with the Howard Hughes Medical Institute. Coordinates will be deposited in the Protein Data Bank (Brookhaven National Laboratory, Upton, New York), and are available before their release by e-mail (ghosh@xtal220.harvard.edu).