Periodic variation in side-chain polarities of T-cell antigenic peptides correlates with their structure and activity

(amphipathicity/hydrophobic moment)

JAMES L. CORNETTE*, HANAH MARGALIT†, JAY A. BERZOFKY‡, AND CHARLES DELISI§

*Department of Mathematics, Iowa State University, Ames, IA 50011; †Department of Molecular Genetics, Hadassah Medical School, Hebrew University, Jerusalem, Israel 91010; ‡Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892; and §Department of Biomedical Engineering, Boston University, Boston, MA 02215

Communicated by David Eisenberg, University of California, Los Angeles, CA, May 5, 1995

ABSTRACT We present an analysis that synthesizes information on the sequence, structure, and motifs of antigenic peptides, which previously appeared to be in conflict. Fourier analysis of T-cell antigenic peptides indicates a periodic variation in amino acid polarities of 3–3.6 residues per period, suggesting an amphipathic α -helical structure. However, the diffraction patterns of major histocompatibility complex (MHC) molecules indicate that their ligands are in an extended non- α -helical conformation. We present two mutually consistent structural explanations for the source of the α -helical periodicity, based on an observation that the side chains of MHC-bound peptides generally partition with hydrophobic (hydrophilic) side chains pointing into (out of) the cleft. First, an analysis of haplotype-dependent peptide motifs indicates that the locations of their defining residues tend to force a period 3–4 variation in hydrophobicity along the peptide sequence, in a manner consistent with the spacing of pockets in the MHC. Second, recent crystallographic determination of the structure of a peptide bound to a class II MHC molecule reveals an extended but regularly twisted peptide with a rotation angle of about 130°. We show that similar structures with rotation angles of 100–130° are energetically acceptable and also span the length of the MHC cleft. These results provide a sound physical chemical and structural basis for the existence of a haplotype-independent antigenic motif which can be particularly important in limiting the search time for antigenic peptides.

A central goal of cell biology is to develop a predictive understanding of the structural basis of cellular activity. With the rapidly increasing size of structural and functional databases, a number of structure–function motifs have been uncovered (e.g., see ref. 1), among them motifs for T-cell antigenic peptides. Findings based on the crystal structure of peptides bound to major histocompatibility complex (MHC) molecules and on sequence motifs for these peptides suggest an interpretation of the periodic variation in side-chain hydrophobicities of T-cell antigenic peptides, providing insights into structure–function relations in antigen presentation.

Cytotoxic/helper T cells recognize antigenic peptides in association with class I/class II products of the MHC, thus stimulating an immune response. Analysis of side-chain hydrophobicity as a function of position for 92 class II antigenic peptides reveals that approximately two-thirds of them show a strong periodic variation, ranging from 3.0 to 3.6 residues per period (2). Similar results are found in several reports (3–8) and for a smaller database of class I-restricted antigenic peptides (2). These findings, which are significant at confidence levels of 10^{-3} and better (2), were interpreted as indicating that (a) period 3.0–3.6 variation in hydrophobicity

values is a dominant, haplotype-independent motif for T-cell antigenicity and (b) the structures of most T-cell antigenic peptides have the potential to form an amphipathic α -helix or a 3_{10} -helix.

Interpretation a found widespread use in algorithms that successfully predicted a number of antigenic sites, including sites in the envelope protein of human immunodeficiency virus (9, 10), the circumsporozoite protein of *Plasmodium falciparum* (11), the envelope protein of human T-lymphotropic virus type I (12), the nucleoprotein (13) and matrix protein (14) of influenza virus, the rat myelin basic protein (15), and the 65-kDa protein of *Mycobacterium tuberculosis* (16). Thus, for both experimental and statistical reasons, the concept of helical amphipathicity seemed to have a reasonable range of validity and applicability—until the class I crystal structure was solved.

Crystal structures of the class I molecule complexed with peptides show an extended rather than an α -helical peptide structure (17–19), indicating that the α -helical interpretation (2) of hydrophobic variation is incorrect. This finding left the relation between hydrophobic variation and T-cell antigenicity without a physical interpretation and necessarily limited its use as a simplifying concept and predictive indicator. We now provide that interpretation in two interrelated ways, and in so doing unify existing data on antigenic motifs and structure. We show that (i) MHC-bound peptides are amphipathic, with the hydrophobic face oriented toward the cleft; (ii) MHC pockets tend to force a 3- to 4-residue periodic variation in side-chain hydrophobicity, similar to that of α -helices; and (iii) the crystallographically observed class II-bound peptide is one of a set of extended, energetically allowable, non- α -helical, amphipathic structures having periodicities in the neighborhood of those characterizing the α -helix. α -Helical amphipathic sequences in native proteins, therefore, tend to be selected as T-cell antigens because the periodicity in the polarity of their side chains resonates with that induced by MHC structure.

METHODS

Databases. The antigenic peptides used to compute Fig. 1 were the 92 used in the analysis Cornette *et al.* (2) and are among the 99 listed by Altuvia *et al.* (20), excluding 7 redundant peptides. The 92 proteins used to select random peptides for Fig. 1 and to search for peptide structure for Fig. 3 were selected, first by the method of Hobohm *et al.* (21) with their most stringent condition that no pair of structures in the database can have >25% homology, then requiring resolution of 2.0 Å or better, and finally excluding some proteins in which several side chains were missing. Protein coordinates were obtained from the Protein Data Bank (22, 23).

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; HA, hemagglutinin.

Simple motifs are found in the sequences of peptides eluted from class I MHC molecules. For example, peptides that bind the murine class I K^b haplotype often have a tyrosine or phenylalanine at position 5 and a leucine or methionine at position 8 or 9, whereas those that bind the human A2.1 haplotype have a leucine or methionine at position 2 and a valine at position 9. MHC allele-specific sequence motifs have also been described for peptides binding to class II molecules, though they are more complex and the limits of their reliability have not been firmly established. A database of 19 class I MHC motifs and 13 human class II motifs appears in ref. 24; 6 murine class II motifs appear in ref. 34, and 6 were assembled from refs. 35–37.

Exposure Hydrophobic Moment. For peptides bound to MHC clefts, the exposure hydrophobic moment is a measure of the tendency for hydrophobic side chains to be exposed to the solvent and hydrophilic side chains to be buried. For each side chain, we compute the ratio (Ratio_{*i*}) of its solvent-accessible surface area when the peptide is in the MHC cleft to the total side-chain surface area when the peptide is outside the cleft. Then the exposure hydrophobic moment is

$$\frac{1}{L} \times \sum_{i=1}^L \left(\frac{\text{Ratio}_i - \text{Ratio}_{\text{ave}}}{\text{Ratio}_{\text{ave}}} \times \text{Hyd}_i \right),$$

where Ratio_{ave} (= 0.259) is the average of the ratios for all non-glycine side chains in the database of nine class I peptides shown in Table 1. We use the hydrophobicity scale of Fauchère and Pliška (29) translated so that the value of glycine is zero and normalized so that the sum of the squares of all values is 25 (30). Hydrophobic and hydrophilic side chains have, respectively, positive and negative hydrophobicities, and the exposure hydrophobic moment is typically negative. When (Ratio_{*i*} – Ratio_{ave})/Ratio_{ave} is positive, the side chain points at least partially up, and multiplication by a positive (hydrophobic) Hyd_{*i*} yields a positive contribution to the moment. When (Ratio_{*i*} – Ratio_{ave})/Ratio_{ave} is negative, the side chain points at least partially down, and multiplication by a positive Hyd_{*i*} yields a negative contribution to the moment. Similarly, upward-pointing polar side chains make negative contributions, and downward-pointing polar side chains make positive contributions.

Power and Least-Squares Spectra. The Fourier power spectrum, $S_p(\theta)$, for the hydrophobicity values of a peptide, p , is

Table 1. Exposure hydrophobic moments (EHM) values

MHC	Antigen*	Ref(s).	PDB file	EHM
HLA-A2	HIV-1 gp120	25	1HHG	–0.57
HLA-A2	HBV nucleocapsid	25	1HHH	–0.44
HLA-A2	Influenza matrix	25	1HHI	–0.67
HLA-A2	HIV-1 RT	25	1HHJ	–0.32
HLA-A2	HTLV-1 Tax	25	1HHK	–0.17
HLA-Aw68	Influenza NP	26	1HSB†	–0.21
HLA-B27	Natural peptide	27	1HSA†	0.48
H2-K ^b	Sendai NP	18	1VAB	–0.61
H2-K ^b	VSV	18,19	1VAA	–0.24
HLA-DR1	Influenza HA	28	†	–0.43

The first nine antigens are class I MHC peptides, eight of which have negative EHM values, indicating that the side chains tend to segregate with the hydrophobic residues pointing into the MHC cleft and hydrophilic residues pointing away from the cleft. The 10th entry, a class II MHC peptide, also has a negative EHM. PDB, Protein Data Bank.

*HIV, human immunodeficiency virus; HBV, hepatitis B virus; RT, reverse transcriptase; HTLV, human T-lymphotropic virus; NP, nucleoprotein; VSV, vesicular stomatitis virus.

†MHC and peptide coordinates were kindly provided by Michael L. Silver (Aw68), Dean R. Madden (B27), and Lawrence J. Stern (DR1).

$$S_p(\theta) = \frac{2}{L} \left\{ \left[\sum_{i=0}^{L-1} (h_i - \bar{h}) \times \cos(i\theta) \right]^2 + \left[\sum_{i=0}^{L-1} (h_i - \bar{h}) \times \sin(i\theta) \right]^2 \right\},$$

where L is the length of p , $\{h_i\}_{i=0}^{L-1}$ is the sequence of hydrophobicity values along the peptide chain, and $\bar{h}$ is the average of the hydrophobicity values. A composite Fourier power spectrum, $S(\theta)$, for a collection of peptides is the sum of the spectra for the individual peptides, $\sum_p S_p(\theta)/L_p$. Subtraction of $\bar{h}$ causes all spectra to be 0 at $\theta = 0$. The spectrum is the hydrophobic moment introduced by Eisenberg *et al.* (31), adjusted by $\bar{h}$, squared, and normalized by L . For a collection of helical peptides, $S(\theta)$ typically has a maximum near 100° which is associated with the 3.6 (360°/100°) residues per turn of a helix and the tendency of helices to be amphipathic (31).

The least-squares spectrum, $T_p(\theta)$, is the sum of squares accounted for by fitting a sinusoidal curve to a sequence of hydrophobicity values and has the basic equation for L odd (30)

$$T_p(\theta) = L\bar{h}^2 + \frac{\left[\sum_{i=-\nu}^{\nu} (h_{i+\nu} - \bar{h}) \cos(i\theta) \right]^2}{\sum_{i=-\nu}^{\nu} \cos^2 i\theta - \frac{1}{L} \left[\sum_{i=-\nu}^{\nu} \cos(i\theta) \right]^2} + \frac{\left[\sum_{i=-\nu}^{\nu} (h_{i+\nu} - \bar{h}) \sin(i\theta) \right]^2}{\sum_{i=-\nu}^{\nu} \sin^2 i\theta},$$

where $\nu = (L - 1)/2$. For $L \geq 15$, say, the least-squares spectrum is essentially indistinguishable from $2/L$ times the square of the hydrophobic moment of Eisenberg *et al.* (31). However, this spectrum is less sensitive to artifacts created by short sequences (30) and is used in the computations of spectra for motifs. The composite least-squares spectrum, $T(\theta)$, for a collection of peptides is $\sum_p T_p(\theta)/L_p$.

An analysis of the binding motifs of class I and class II MHC molecules was made by using a composite least-squares spectrum. An example motif for HLA-A2 (24) is as follows.

	P ₁	P ₂	P ₃	P ₄	P ₅	P ₆	P ₇	P ₈	P ₉
Anchor		L							V
Strong		M		E,K		V		K	

For such a motif, a sequence of hydrophobicity values was constructed by using the average hydrophobicity values of the dominant side chains at each position, with 0.0 (glycine) for the empty positions. The dominant side chains and hydrophobicity values for the HLA-A2 motif are as follows.

P ₁	P ₂	P ₃	P ₄	P ₅	P ₆	P ₇	P ₈	P ₉
G	L	G	E,K	G	V	G	K	V
0.0	1.70	0.0	–0.81	0.0	1.22	0.0	–0.99	1.22

Extended but Twisted Conformations. The space of main-chain ϕ , ψ angles that yield extended but twisted conformations similar to that of the influenza hemagglutinin (HA)–(307–319) peptide bound in the class II MHC molecule DR1 (28) was examined as follows. Polyalanine 13-mers were constructed with standard geometry and integer values of $-180^\circ \leq \phi \leq -60^\circ$ and $60^\circ \leq \psi \leq 180^\circ$. An axis was fit to the polymer by a method described previously (30), and the twist of the peptide about that axis was computed. The 13-mers

having C_1^α – C_{13}^α lengths within 2.0 Å of that of the HA-(307–319) peptide (37.2 Å) trace out in the ϕ – ψ plane a crescent-shaped swath whose central axis represents structures 37.2 Å in length. The energies of 16 of these latter structures, chosen at uniform intervals, were evaluated by use of CHARMM 21 (32). Without minimization, all peptides except one (at $\phi = -180^\circ$) had negative van der Waals energies and total energies between -157 kcal/mol and -253 kcal/mol. We conclude that the peptides with constant ϕ – ψ values in the crescent region are sterically acceptable.

We searched the 92 proteins in our structural database for backbone conformations resembling that of the HA peptide by using a least root-mean-square (rms) fit of the HA peptide backbone to every 13-mer peptide in the database. Because of a slight reversal in the HA peptide twist between residues 3 and 4, we also searched for peptides that were conformationally close to the backbone of the C-terminal 10 residues of the HA peptide.

RESULTS AND DISCUSSION

A composite Fourier spectrum is shown in Fig. 1 for 92 class II antigenic peptides, together with the spectra for five sets of randomly selected protein segments. The peak at 96° in the spectrum for antigenic peptides, as distinct from that of randomly selected peptides, provides a simple graphic indication of a nonrandom periodic repetition of hydrophobicity values along the antigenic peptide sequences of $360^\circ/96^\circ \approx 3.7$ residues per period. Several analyses have confirmed the statistical significance of the periodic repetition of hydrophobicity values (2–8), and the repetition itself has been useful as an indicator of potential T-cell antigenic sites (9–16).

The peak in the spectrum for antigenic peptides is similar to the peak at 100° in the spectrum for α -helical peptides found by Eisenberg *et al.* (31), and one might expect the structure of antigenic peptides to be α -helical when bound to MHC molecules. However, the 10 known structures of peptides embedded in MHC clefts are all clearly extended and nonhe-

lical. Other physical properties are therefore required to interpret the observed periodic repetition. We provide two interrelated interpretations that unify existing data on antigenic motifs and structure.

Amphipathic Separation of Peptide Side Chains. Analysis of the exposure hydrophobic moments of MHC-bound peptides indicates that they are amphipathic, with 9 of the 10 having their hydrophobic face oriented toward the MHC binding cleft (Table 1). The peptide backbones generally lie parallel to the interface between hydrophobic and hydrophilic regions in an extended conformation. Any structural constraint, such as pockets in the MHC cleft or a twist in the peptide about its backbone, that causes peptide side chains to extend into one of the regions at regularly spaced intervals will be reflected in periodic repetition of hydrophobicity values along the peptide primary sequence.

Periodicity in Motifs. The first interpretation of periodic variation in polarity of T-cell antigens stems from simple haplotype-dependent motifs found in the sequences of peptides eluted from class I MHC molecules, and from similar though less distinct motifs for class II MHC molecules. These motifs, though neither necessary nor sufficient for antigenicity, are nevertheless extremely valuable for identifying protein segments that are antigenic for a particular haplotype. Dominating each motif are the anchor residues that bind in pockets in the MHC cleft; nonanchor residues typically point out of or across the cleft at peptide positions not corresponding to pockets.

Least-squares spectra of the motifs for the class I and class II antigenic peptides (Fig. 2) demonstrate a periodicity in side-chain polarity of the motif residues alone. Because of spaces between motif residues that are filled with glycines in our analysis, there is an artifactually dominant frequency at 0° in both cases. However, a sharp secondary peak at 93° can be seen in the spectrum of the class I motifs. Even though there were only 19 class I motifs as opposed to 25 class II motifs, the greater definition of the class I motifs is apparent from the sharper secondary peak in the class I spectrum. The secondary

FIG. 1. Periodic variation in side-chain polarities of antigenic peptides. A composite Fourier power spectrum is shown for 92 class II T-cell antigenic peptides (solid line) and five sets of 92 peptides of corresponding lengths randomly selected from a collection of 92 nonhomologous proteins (dotted and dashed lines). The peak at 96° of the spectrum for antigenic peptides, as distinct from the spectra of randomly selected peptides, indicates a nonrandom periodic repetition of hydrophobicity values along the peptide sequences of $360^\circ/96^\circ \approx 3.7$ residues per period.

FIG. 2. Composite least-squares spectra, $T(\theta)$, of 19 binding motifs of class I (solid curve) and 25 binding motifs of class II (dotted curve) MHC molecules. The spectra were normalized by dividing by 1/10th the number of motifs. With the large blocks of zero entries in each motif, the artifactually dominant frequency is 0 (a constant function is best) in both cases. However, each set of motifs shows a more meaningful dominant secondary peak, at 93° (period 4) for class I and at 122° (period 3) for class II.

peak in the class II spectrum occurs at 122° . The composite spectrum of the motifs over various haplotypes indicates that the locations of the predominantly hydrophobic anchor residues force a period 3.0–4.0 variation in hydrophobicity values on the peptides, which is also consistent with the spacing of hydrophobic pockets in the MHC molecule. It therefore seems plausible that the reason α -helical amphipathic structures tend to be selected as T-cell antigens is that they have a periodicity similar to that induced by pockets of the MHC cleft.

Twist in the HA-(30–319) Peptide. A second, supporting, and very striking interpretation of period 3–4 variation in side-chain hydrophobicity found in antigenic peptides is provided by the recently solved class II DR1 structure (28), whose bound ligand, the influenza HA-(307–319) peptide, is highly extended but at the same time has a left-handed twist of -130° with predominantly hydrophilic side chains (Lys, Val, Lys, Asn, Lys, and Thr) pointing out ($\uparrow$) of the binding cleft and predominantly hydrophobic side chains (Pro, Tyr, Gln, Thr, Leu, Leu, and Ala) pointing toward or across ($\bullet$) the cleft: $\bullet \uparrow \bullet \uparrow \bullet \uparrow \bullet \uparrow \bullet \uparrow \bullet \uparrow \bullet \uparrow \bullet \uparrow$. The spatial separation of hydrophobic and hydrophilic residues together with approximate period 3 rotation (120° rotation) of the side chains along the peptide is reflected in a period 3 variation of hydrophobicity values along the sequence. None of the nine crystallized class I peptides exhibit such twisting.

The set of polyaniline tridecamers having lengths within 2 Å of the HA peptide cuts a crescent swath in the ϕ - ψ plane which encompasses structures ranging from those that have a left-handed twist of -100° per residue to a right-handed twist of 120° per residue (Fig. 3A). The backbone structure of a single peptide used in the computation of Fig. 3A is shown in Fig. 3B, where it can be seen that consistent $\phi = -80^\circ$, $\psi = 148^\circ$ torsion angles cause the peptide backbone to spiral around an axis.

Four- and 5-residue peptides having the left-handed twist of the HA peptide are common in globular proteins (33), but 10-residue and longer peptides of that conformation are rare. A search of 92 protein structures for a backbone conformation resembling that of the HA peptide yielded none with rms distance less than 1.4 Å. The C-terminal decamer of HA is a more uniformly twisted part of HA ($\phi = -88^\circ \pm 16^\circ$, $\psi = 144^\circ \pm 16^\circ$), and there were 27 peptides with a backbone rms distance less than 1.4 Å from the C-terminal decamer. Twenty-three of them contained 4 or more β -strand residues. Visual inspection of these peptides showed a left-handed twist in at least part of each peptide; only one (8ACN 539–548, rms distance = 1.06) showed a left-handed twist along the whole peptide. The HA C-terminal decamer is remarkably uniform in its ϕ , ψ values, with 16° standard deviation in each value. Only 3 of the 27 peptides found here had ϕ , ψ standard deviations less than 20° in each component (2MNR 31–40, $-106^\circ \pm 19^\circ$, $132^\circ \pm 18^\circ$; 8ACN 539–548, $-74^\circ \pm 13^\circ$, $143^\circ \pm 10^\circ$; 8ACN 709–818, $-91^\circ \pm 18^\circ$, $132^\circ \pm 17^\circ$).

The extended but twisted structure of the HA peptide rarely occurs in the database but belongs to a set of structures that are sterically acceptable over a range of twist values. In addition, a large range of ϕ , ψ values yield peptides with MHC cleft-compatible dimensions (Fig. 24). Although the structures are uncommon in native proteins and would not be expected to be stable in solution, interactions with the MHC groove provide the energy to stabilize them when bound to the MHC molecule. The hydrophobic periodicity is enforced by the spacing of the MHC pockets which bind anchor residues of the peptide. As more class II complexes are crystallized, it would not be surprising to find other peptides falling within the crescent region.

Taken collectively, these results suggest that extended structures can be associated with various periodicities in side-chain

FIG. 3. Extended but twisted peptide conformations. (A) Regions of the upper left corner of the Ramachandran ϕ - ψ map are shown for which a 13-mer peptide with constant ϕ - ψ values yields an extended but twisted peptide of C_α - C_α length close to that of the 13-mer HA peptide (37.2 Å). The crescent-shaped region yields peptides whose geometric lengths are within 2.0 Å of that of the HA peptide, and the central curve locates points that give the HA peptide length (37.2 Å) exactly. The lines marked 3.6, 3.0, and 2.0 show the number of residues per turn of the peptides. Each white dot marks the average ϕ , ψ values of a 10-residue segment from the database of 92 structures for which the backbone atoms could be fit to the backbone of the C-terminal 10-mer of the HA antigenic peptide with rms distance of all backbone atoms ≤ 1.4 Å. The black square marks the average ϕ , ψ values (-90° , 145°) of the HA peptide and the black circle marks the ϕ , ψ values (-78° , 149°) of the type II polyproline helix. (B) View from the N terminus and side view of a peptide with constant $\phi = -80^\circ$ and $\psi = 148^\circ$ values, yielding a left-handed twist of -122° . The size of the circle at an atom position indicates distance from the viewer. In the side view, the lighter shade indicates atoms that are on the viewer's side of the page (and are to the right in the N-terminal view), and the dashed line is the central axis. A vertical separation of hydrophobic and hydrophilic peptides in this structure would lead to period 3 variation in the polarities of the side chains along the primary sequence.

properties (i.e., a necessary association of 100° periodicity with an α -helical or any other particular structure is not correct) and that periodic variation in hydrophobicity values remains a strong indicator of antigenicity. The natural tendency to associate particular periodicities with prevalent structural types such as the α -helix must be expanded to include the extended-twisted structural type found in the HA peptide. That structure, together with the observation that, generally, hydrophobic side chains point into the cleft and hydrophilic side chains point out, leads to periodic variation in hydrophobicity values along the peptide.

Conclusions. The peptide backbone twist observed in the class II crystal structure and the analysis of peptide motifs lead to an unexpected explanation of the statistical correlation observed earlier between helper-T-cell antigenic sites and the periodic variation in hydrophobicity along the sequence. Although individual motifs are specific for individual MHC loci and alleles, the analysis shown here demonstrates a more general supermotif that is common to most MHC alleles but, like individual motifs, enforced by the spacing and character of the MHC pockets into which anchor residues must fit. This supermotif is compatible with the known crystal structures of peptide-MHC complexes. Thus, the periodic variation in side-chain properties, indicative of amphipathicity, remains an important indicator of antigenicity.

We thank Zhiping Weng for computation of peptide side-chain exposures and Benjamin King for assistance with graphics. This work was supported in part by Grant AI30535 from the National Institutes of Health.

- Marshall, G. (1992) *Curr. Opin. Struct. Biol.* **2**, 904–919.
- Cornette, J. L., Margalit, H., DeLisi, C. & Berzofsky, J. A. (1993) in *The Amphipathic Helix*, ed. Epand, R. M. (CRC, Boca Raton, FL), pp. 333–345.
- DeLisi, C. & Berzofsky, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7048–7052.
- Margalit, H., Spouge, J. L., Cornette, J. L., Cease, K. B., DeLisi, C. & Berzofsky, J. A. (1987) *J. Immunol.* **128**, 2213–2229.
- Spouge, J. L., Cornette, J. L., Margalit, H., Cease, K., Guy, H. R., Berzofsky, J. A. & DeLisi, C. (1987) *J. Immunol.* **138**, 204–212.
- Stille, C. J., Thomas, L. J., Reys, V. E. & Humphreys, R. E. (1987) *Mol. Immunol.* **24**, 1021–1027.
- Rothbard, J. B. & Taylor, W. R. (1988) *EMBO J.* **7**, 93–100.
- Reyes, V. E., Fowlie, E. J., Lu, S., Phillips, L., Chin, L. T., Humphreys, R. E. & Lew, R. A. (1990) *Mol. Immunol.* **27**, 1021–1027.
- Cease, K. B., Margalit, H., Cornette, J. L., Putney, S. D., Robey, W. G., Ouyang, C., Streicher, H. Z., Fischinger, P. J., Gallo, R. C., DeLisi, C. & Berzofsky, J. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4249–4253.
- Takahashi, H., Cohen, J., Hosmalin, A., Cease, K. B., Houghten, R., Cornette, J. L., DeLisi, C., Moss, B., Germain, R. N. & Berzofsky, J. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3105–3109.
- Good, M. F., Maloy, W. L., Lunde, M. N., Margalit, H., Cornette, J. L., Smith, G. L., Moss, B., Miller, L. H. & Berzofsky, J. A. (1987) *Science* **235**, 1059–1062.
- Kurata, A., Palker, T. J., Streilein, R. D., Searce, R. M., Haynes, B. F. & Berzofsky, J. A. (1989) *J. Immunol.* **143**, 2024–2030.
- Bastin, J., Rothbard, J., Davey, J., Jones, I. & Townsend, A. (1987) *J. Exp. Med.* **155**, 1508–1523.
- Gotch, F., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. (1987) *Nature (London)* **326**, 881–882.
- Zamvil, S., Mitchell, D., Moore, A., Kitamura, K., Steinman, L. & Rothbard, J. (1986) *Nature (London)* **324**, 258–260.
- Lamb, J., Ivanyi, J., Rees, A., Rothbard, J., Howland, K., Young, R. & Young, D. (1987) *EMBO J.* **6**, 1245–1249.
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1991) *Nature (London)* **353**, 321–325.
- Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257**, 919–927.
- Zhang, W., Young, A. C. M., Imarai, M., Nathenson, S. G. & Sacchettini, J. C. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8403–8407.
- Altuvia, Y., Berzofsky, J. A., Rosenfeld, R. & Margalit, H. (1994) *Mol. Immunol.* **31**, 1–19.
- Hobohm, U., Scharf, M., Schneider, R. & Sander, C. (1992) *Protein Sci.* **1**, 409–417.
- Bernstein, F. C., Koetzle, T. F., Williams, B. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542.
- Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. (1987) in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*, eds. Allen, F. H., Bergerhoff, G. & Sievers, R. (Data Comm. Int. Union Crystallogr., Bonn), pp. 107–132.
- Meister, G. E., Roberts, C. G. P., Berzofsky, J. A. & De Groot, A. S. (1995) *Vaccine* **13**, 581–591.
- Madden, D. R., Gargoczi, D. N. & Wiley, D. C. (1993) *Cell* **75**, 693–708.
- Silver, M. L., Guo, H.-C., Strominger, J. L. & Wiley, D. C. (1992) *Nature (London)* **360**, 367–369.
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1992) *Cell* **70**, 1035–1048.
- Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L. & Wiley, D. C. (1994) *Nature (London)* **368**, 215–221.
- Fauchère, J. L. & Pliška, V. (1983) *Eur. J. Med. Chem.* **18**, 369–375.
- Cornette, J. L., Cease, K. B., Margalit, H., Spouge, J. L., Berzofsky, J. A. & DeLisi, C. (1987) *J. Mol. Biol.* **195**, 659–685.
- Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 140–144.
- Brooks, C. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4**, 187–213.
- Adzhubei, A. A. & Sternberg, M. J. E. (1993) *J. Mol. Biol.* **229**, 472–493.
- Rammensee, H.-G., Friede, T. & Stevanović, S. (1995) *Immunogenetics* **41**, 178–228.
- Corr, M., Boyd, L. F., Padlan, E. A. & Margulies, D. H. (1993) *J. Exp. Med.* **178**, 1877–1892.
- Sette, A., Buss, S., Appella, E., Smith, J. A., Chesnut, R., Miles, C., Colon, S. M. & Grey, H. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3296–3300.
- Leighton, J., Sette, A., Sidney, J., Appella, E., Ehrhardt, C., Fuchs, S. & Adorini, L. (1991) *J. Immunol.* **147**, 198–204.