Biochemistry **1994**, *33*, 2055–2062

NMR-Derived Solution Conformations of a Hybrid Synthetic Peptide Containing Multiple Epitopes of Envelope Protein gp120 from the RF Strain of Human Immunodeficiency Virus[†]

Robert de Lorimier,[‡] M. Anthony Moody,[§] Barton F. Haynes,[§] and Leonard D. Spicer*,[‡]

Departments of Biochemistry, Radiology, Immunology, and Medicine and Duke Center for AIDS Research, Duke University Medical Center, Durham, North Carolina 27710

Received July 20, 1993; Revised Manuscript Received December 6, 1993[●]

ABSTRACT: Solution conformations of a 40-residue hybrid peptide containing T-helper epitopes and B-cell determinants from envelope glycoprotein gp120 of human immunodeficiency virus (HIV) have been investigated with NMR. Peptides of this general design are highly immunogenic and induce HIV-neutralizing antibodies and T-lymphocyte responses. The 16-residue N-terminal segment of the peptide contains a T-helper epitope, while the 24-residue C-terminal segment is derived from the V3 loop of HIV strain RF and contains epitopes that elicit neutralizing antibodies as well as T-cell responses. On the basis of 2D proton NMR spectra (COSY, TOCSY, and NOESY) of the peptide in aqueous solution, the resonances of nearly all hydrogens are assigned. The peptide is largely disordered, but specific medium-range NOEs demonstrate conformational preferences in certain regions. Part of the N-terminal segment exhibits nascent helical conformations, consistent with a finding that many T-cell antigens can be modeled as amphipathic helices. In the V3-derived segment of the peptide, one region shows evidence of a tight turn conformation, corresponding to a turn found previously in V3 peptides of HIV strains MN and III$_B$. Other conformational features are also detected in the V3 region, such as a stretch of β strand and a kink that may arise from side-chain interactions.

The observation that peptide fragments derived from proteins sometimes induce antibodies which bind to the intact native protein has stimulated increasing interest in the relationship between peptide conformational preferences and functional activities such as binding and immunogenicity. Free peptides in solution generally do not exhibit the stable secondary structure characteristic of folded proteins, but they do show nascent structural organization, detectable by NMR or circular dichroism, which can be similar to the corresponding domain in the intact protein or to the structure assumed by the peptide in a bound state. For example, a peptide fragment of the protein hirudin binds to thrombin in a conformation which resembles preferentially populated conformers of the peptide free in solution (Ni et al., 1992). Similarly, the conformation of an immunogenic peptide from myohemerythrin has been examined both free in solution (Dyson et al., 1988b) and when bound to the F$_{ab}$ portion of an antibody that was induced by the peptide and which binds native myohemerythrin (Tsang et al., 1992). It was found that the main-chain conformation over part of the free peptide resembled its conformation in the F$_{ab}$-bound complex. Such observations are consistent with a model suggesting that peptide conformation is important in binding to proteins.

The property of immunogenicity in a peptide has been hypothesized to be due to the presence of preferred conformers which determine binding to receptors on B lymphocytes (Lerner, 1984; Dyson et al., 1988a). To explore this hypothesis, several studies have examined immunogenic peptides for propensities to adopt preferential conformations. Such propensities have been observed in immunogenic peptides derived from hemagglutinin of influenza virus (Dyson et al., 1985), 24M protein of *Streptococcus* (Gras-Masse et al., 1988), gD-1 protein of herpes simplex virus (Williamson et al., 1986), coat protein from *Plasmodium* (Dyson et al., 1990), envelope glycoprotein gp120 from human immunodeficiency virus (Chandrasekhar et al., 1991; Zvi et al., 1992), and transmembrane protein from simian immunodeficiency virus (Dyson et al., 1992a).

Peptide conformation could also play a role in another immune response, the activation of T lymphocytes. The antigen recognized by T cells is not a native protein but peptide fragments bound to major histocompatibility complex (MHC)[1] molecules on the surface of antigen-presenting cells. The protein antigen is rendered in this form by intracellular processing events (Germain & Margulies, 1993). Hence, for an exogenous immunogenic peptide to elicit a T-cell response, it must become bound to MHC molecules in a manner that mimics the naturally processed peptide fragments. The determinants of peptide antigenicity for T cells are not fully defined, but some correlates have been described, including linear sequence (Falk et al., 1991; Rothbard & Taylor, 1988) and secondary structure motifs (DeLisi & Berzofsky, 1985; Schrier et al., 1989; Spouge et al., 1987). These motifs may very well be related not only to antigen recognition by T cells

---

[†] Supported in part by grants from the National Institutes of Health (GM 41829 and RR 05748 to L.D.S. and CA 43447 and AI 28662 to B.F.H.), the National Science Foundation (9011410 to L.D.S.), and the North Carolina Biotechnology Center (9010-1DG-1004 to L.D.S.). R.d.L. was supported by NIH Training Grant A107217. B.F.H. is a Carter Wallace Fellow in Retroviral Research.

[‡] Departments of Biochemistry and Radiology.

[§] Departments of Immunology and Medicine and Duke Center for AIDS Research.

[●] Abstract published in *Advance ACS Abstracts*, February 1, 1994.

[1] Abbreviations: HIV, human immunodeficiency virus; gp120, glycoprotein of 120-kDa apparent mass; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser and exchange spectroscopy; NOE, nuclear Overhauser effect; Th, T helper (lymphocyte); CTL, cytotoxic T lymphocyte; DTT, dithiothreitol; MHC, major histocompatibility complex; $dxy(i,i+n)$, connectivity between a hydrogen attached to atom x on residue i and a hydrogen attached to atom y on residue $i+n$, where n = 1, 2, 3, ...

0006-2960/94/0433-2055$04.50/0

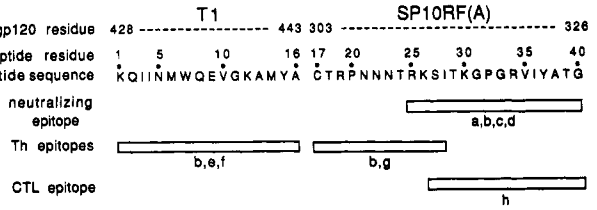

FIGURE 1: Derivation of peptide T1SP10RF(A) from HIV envelope glycoprotein gp120. Numbering of gp120 corresponds to the *env* gene product of strain III_B (Ratner et al., 1985). Bars delineate regions that are immunogenic for neutralizing antibodies, T-helper (Th) lymphocytes, or cytotoxic T lymphocytes (CTL), as indicated on the left. Letters below bars indicate the following references: (a) Palker et al., 1988; (b) Palker et al., 1989b; (c) Rusche et al., 1988; (d) Goudsmit et al., 1988; (e) Cease et al., 1987; (f) Wahren et al., 1989; (g) Schrier et al., 1989; (h) Takahashi et al., 1992.

but also to the numerous steps in processing the antigen for presentation, such as proteolysis, transport into the endoplasmic reticulum, and assembly with MHC protein. Thus, the conformation of a peptide fragment in solution can be a factor in each of these steps. In only a few cases has the solution conformation of T-cell antigenic peptides been examined (Abergel et al., 1989; Lark et al., 1989; Waltho et al., 1989). Such studies should be useful in determining whether conformation is a determinant of MHC-associated antigenicity.

We are using proton NMR to investigate the conformations of hybrid synthetic peptides that contain B-cell and T-cell epitopes from noncontiguous regions of the gp120 envelope protein of human immunodeficiency virus (HIV) (Palker et al., 1989b). In animal immunization studies, these peptides elicit (a) neutralizing antibodies to native gp120 (Palker et al., 1988, 1989a,b; Hart et al., 1990; Haynes et al., 1993), (b) T-helper (Th) cells that proliferate when exposed to target cells (Palker et al., 1989b; Hart et al., 1990), and (c) cytotoxic T lymphocytes (CTL) that lyse target cells expressing native HIV gp120 (Hart et al., 1991). Hybrid peptides containing covalently linked B- and T-cell epitopes of gp120 are considerably more immunogenic for both B- and T-cell responses than peptides containing single separate epitopes (Palker et al., 1989a,b; Shirai et al., 1994).

The peptide that is the subject of this report is designated T1SP10RF(A) and is derived from gp120 of HIV strain RF (Starcich et al., 1986). Its sequence is similar to a previously reported hybrid peptide (Palker et al., 1989b), but bears an additional three residues (Ala-Thr-Gly) at the carboxyl terminus. Like the previously reported peptide, T1SP10RF(A) induces high titers of antibodies to HIV strain RF (T. J. Matthews, A. M. Moody, and B. F. Haynes, unpublished results). The N-terminal 16 residues of T1SP10RF(A) (Figure 1) occur at positions 428–443 of gp120 and comprise a potent Th determinant called T1 (Cease et al., 1987; Schrier et al., 1989; Dadaglio et al., 1991) that is thought to be part of the CD4 binding site of gp120 (Lasky et al., 1987; Sun et al., 1989). The segment of the hybrid peptide from residues 17 to 40 is called the SP10RF(A) region (Figure 1) and is equivalent to gp120 residues 301–324, which occur in the third variable (V3) region (Starcich et al., 1986). These peptide residues elicit neutralizing antibodies (Palker et al., 1988, 1989b; Rusche et al., 1988; Goudsmit et al., 1988), Th responses (Palker et al., 1989b; Schrier et al., 1989; Wahren et al., 1989), and MHC class I-restricted CTL responses (Hart et al., 1991; Takahashi et al., 1988, 1989, 1992).

The numerous and relatively well-defined epitopes of T1SP10RF(A) make it an attractive subject for NMR spectroscopic studies of solution conformation in an immunogenic peptide. The information derived from these studies can be combined with structural and immunological data on individual epitopes and on related peptides from other strains of HIV to gain inferences on the role of conformation in eliciting T- and B-cell responses.

MATERIALS AND METHODS

Peptide Synthesis and Characterization. Peptide T1SP10RF(A) was synthesized by fluorenylmethoxycarbonyl chemistry on an ABI 431A peptide synthesizer (Applied Biosystems, Inc., Foster City, CA). Purification was by high-performance liquid chromatography (HPLC). Purity was confirmed by HPLC, amino acid analysis, and electrospray mass spectrometry, as described previously (Haynes et al., 1993). Mass spectrometry was performed by Dr. Robert D. Stevens, Department of Pediatrics, Duke University Medical Center. The calculated mass of the peptide was 4497.26 Da; the observed mass was 4497.14 Da.

Peptide Solutions. It was assumed that the lyophilized solid peptide stock contained no water. Three solution conditions were used: (a) 2 mM peptide, pH 4.4; (b) 4 mM peptide, 20 mM NaCl, and 5 mM KH_2PO_4, pH 3.8; (c) 4 mM peptide, 20 mM NaCl, 10 mM dithiothreitol (DTT), and 5 mM KH_2PO_4, pH 4.2. All solutions contained 90% ^1H_2O, 10% ^2H_2O, 1 mM sodium azide, and 0.5 mM sodium 3-(trimethylsilyl)propionate. The methyl hydrogen resonance of the latter served as a chemical shift reference (0.00 ppm). The pH of solutions was adjusted by adding small volumes of 0.1 N NaOH or 0.1 N HCl.

NMR Spectroscopy. Proton NMR spectra were collected on a Varian Unity 500-MHz spectrometer (static field 11.75 T) at temperatures from 278 to 303 K. Deuterium in the sample (^2H_2O) provided the frequency lock signal. Two-dimensional spectra were obtained with the following pulse sequences: (a) double-quantum-filtered correlation spectroscopy (DQF-COSY) (Piantini et al., 1982; Rance et al., 1983); (b) total correlation spectroscopy (TOCSY) (Bax & Davis, 1985; Levitt et al., 1982) with a mixing time of 170 ms; and (c) nuclear Overhauser effect spectroscopy (NOESY) (Jeener et al., 1979) with mixing times ranging from 75 to 300 ms. Water resonance was suppressed by selective saturation during the recycling delay and, for NOESY, during the mixing period. The spectral width was 6700 Hz, with the indirectly acquired dimension collected as 256 complex increments and the directly acquired dimension containing 1024 complex points. Data were processed with FELIX 1.0 software (Hare Research Inc.). Free-induction decays in the directly acquired dimension were corrected for base-line offset. Decays in both dimensions were multiplied by a sinebell-squared function (phase-shifted by 60 to 90°) and zero-filled to 2048 points prior to Fourier-transformation.

Temperature-Shift Coefficients. TOCSY spectra (mixing time of 50 ms) were acquired at seven temperatures from 278 to 296 K in increments of 3 K. The chemical shifts of 27 backbone amide hydrogens were determined unambiguously over this range and used to calculate, by linear least-squares fitting, the dependence of chemical shift on temperature.

RESULTS

Resonance Assignment. Amino acid residues of peptide T1SP10RF(A) were typed by DQF-COSY, TOCSY, and NOESY, and assigned to specific positions by NOESY using common strategies (Wüthrich, 1986). Overlap of intraresidue NOESY cross-peaks between amide hydrogens and α-carbon hydrogens, $dN\alpha$, with interresidue $d\alpha N(i,i+1)$ NOESY cross-

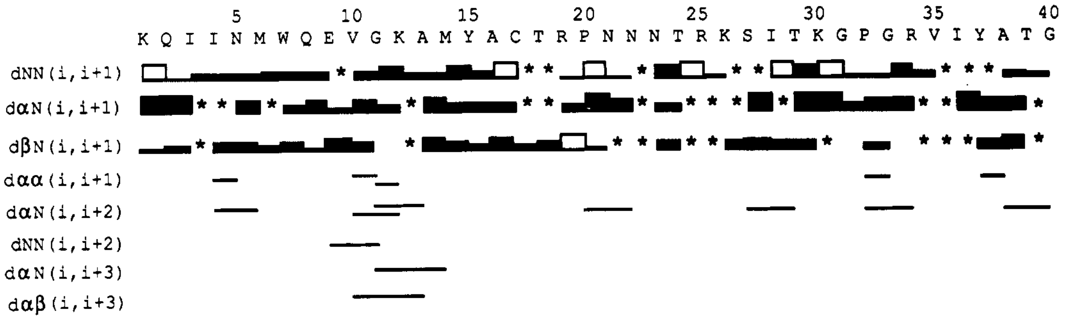

FIGURE 2: Interresidue NOE connectivities in T1SP10RF(A). In rows dNN(i,i+1), $d\alpha$N(i,i+1), and $d\beta$N(i,i+1), the thickness of the solid boxes is proportional to the intensity (peak volume) of the NOE cross-peak. An asterisk represents an NOE of unknown intensity due to overlap with another peak. A blank box indicates an undetected NOE. In rows $d\alpha\alpha$(i,i+1), $d\alpha$N(i,i+2), dNN(i,i+2), $d\alpha$N(i,i+3), and $d\alpha\beta$(i,i+3), the lines connect residues showing an unambiguous NOE. Connectivities to NH of prolines actually refer to C$^{\delta}$H.

peaks caused ambiguities in the sequential assignment at some positions, but these were resolved by the presence of dNN(i,i+1) and/or $d\beta$N(i,i+1) NOESY connectivities (Figure 2). Comparison of spectra obtained at 278 and 303 K also aided the resolution of ambiguities in the assignment. As a result, nearly all nonexchangeable and slowly exchanging hydrogens were assigned. Table 1 shows the chemical shift values of the assigned hydrogens at 278 K in the pH 4.2 buffer containing 10 mM DTT. In this solution, the peptide was entirely monomeric, as opposed to its state in buffer lacking DTT where the peptide was largely dimerized (see below). Except where noted, the following discussion pertains to peptide at 278 K in solution containing DTT. To test if any spectral features were concentration-dependent, a one-dimensional spectrum was acquired of peptide at 0.1 mM in buffer containing DTT. Comparison of this spectrum with that of peptide at 4 mM showed no detectable differences in chemical shift values, and line widths were identical within 1 Hz.

Overall Peptide Characteristics. On the basis of the NMR data, the average configuration of peptide T1SP10RF(A) appears to be highly disordered in solution. The chemical shifts of almost all hydrogens are close to their values in model random-coil peptides (Wüthrich, 1986). This general disorder is illustrated in the chemical shifts for C$^{\alpha}$ hydrogens, which in proteins have been correlated with regular secondary structure (Wishart et al., 1991). Figure 3 shows the amino acid specific C$^{\alpha}$H shift difference between T1SP10RF(A) and the mean for several proteins. On the basis of the criteria of Wishart et al., there is little evidence in these data for stable, regular secondary structure. The pattern of interresidue NOEs between backbone hydrogens of the peptide (Figure 2) also suggests the peptide explores multiple conformations. For example, both dNN(i,i+1) and $d\alpha$N(i,i+1) NOEs are present, which is expected for peptides fluctuating between different backbone configurations (Dyson & Wright, 1991). In addition, there are few medium-range interresidue NOEs, and longer range NOEs (greater than four residues) are not observed. However, certain segments of the peptide do show NOEs, indicating that particular conformations are frequently attained, as discussed below.

Conformations in the T1 Segment. The N-terminal three residues of the T1 region appear to be in an extended conformation as judged by the lack of medium-range interresidue connectivities and the presence of strong $d\alpha$N(i,i+1) NOEs. The stretch from Ile^{4} to Ala^{16} shows NOEs which, considered collectively, are characteristic of nascent helical conformations. As diagrammed in Figure 2, there is an uninterrupted stretch of dNN(i,i+1) NOEs (illustrated in Figure 4) along with relatively weak $d\alpha$N(i,i+1) NOEs, suggestive of tandem backbone ϕ and φ dihedral angles found in the α region of the ϕ,φ plane. The regions from Val^{10} to Ala^{13} and from Ile^{4} to Met^{6} show $d\alpha$N(i,i+2) NOEs which are indicative of nascent helix formation, as is the dNN(i,i+2) NOE between Glu^{9} and Gly^{11} (Figure 4) (Dyson et al., 1988b). In particular, the $d\alpha$N(i,i+3) and $d\alpha\beta$(i,i+3) NOEs (Figure 5) in the region from Val^{10} to Met^{14} are highly indicative of a full helical turn (Wüthrich, 1986; Dyson & Wright, 1991). One residue near this segment, Gln^{8}, shows an anomalously low temperature-shift coefficient ($-\Delta\delta/\Delta T \times 10^{3}$ ppm/K) for its backbone amide hydrogen. The value of this coefficient for Gln^{8} is 2.9×10^{-3} ppm/K, whereas for all other measured amide hydrogens it is greater than 5.7×10^{-3} ppm/K [mean = $(7.6 \pm 1.0) \times 10^{-3}$ ppm/K, n = 27]. Such a low value for Gln^{8} is suggestive of hydrogen bond formation (Wright et al., 1988). The proximity of this residue to a region that adopts a helical turn suggests that the Gln^{8} backbone amide hydrogen is involved in the formation of helical structures. However other mechanisms, such as hydrogen bonding to the carboxylate of Glu^{9}, could also result in a low temperature-shift coefficient.

Conformations in the V3 [SP10RF(A)] Segment. A distinctive conformation appears in the sequence from residues 30 to 34 (Lys-Gly-Pro-Gly-Arg). There is a pattern of NOEs (Figure 2) indicative of a hairpin turn, wherein residues Gly^{31}–Arg^{34} comprise positions 1–4, respectively, of the putative turn. A weak $d\delta$N(i,i+1) NOE between Pro^{32} and Gly^{33}, as well as a strong dNN(i,i+1) NOE between Gly^{33} and Arg^{34}, fits a pattern expected for residues 2, 3, and 4 of a β turn (Dyson et al., 1988c). A $d\alpha$N(i,i+2) NOE between turn residues 2 and 4 (Pro^{32} and Arg^{34}) is expected for a β turn, and is observed (Figure 6). The relatively weak $d\alpha$N(i,i+1) NOE between Pro^{32} and Gly^{33} suggests conformations resembling a type I turn, as does the detectable $d\delta$N(i,i+1) NOE between these residues. Two other features are consistent with a tight turn involving residues 30–34. A dNN(i,i+3) NOE between residues 1 and 4 is a frequent feature of turns connecting antiparallel β sheets in structured proteins (Wüthrich, 1986), and is found for Gly^{31} and Arg^{34} (Figure 4, cross-peak at 8.25/8.46 ppm). An unambiguous identification of this NOE is prevented by the occurrence of other residues having similar resonances. However, all other options accounting for this NOE represent longer range interactions, and they would occur in regions lacking evidence of structure which could account for them. A second feature in the region from Lys^{30} to Arg^{34} is the lack of a dNN(i,i+1) NOE, but the presence of a strong $d\alpha$N(i,i+1) NOE, between Lys^{30} and Gly^{31}. This suggests a strong propensity for ϕ and φ dihedral angels of Lys^{30} that are typical of the β region of ϕ,φ space. Taken together, these features are evidence for conformations resembling a β turn in this portion of T1SP10RF(A).

Table 1: ^{1}H Resonance Assignments of T1SP10RF(A)

residue	chemical shift (ppm)^a NH	αH	βH	other side-chain H
Lys^1	nd	4.02	1.90	γH, 1.43; δH, 1.70; ϵH, 3.00; NH, 7.65
Gln^2	8.91	4.40	1.97, 2.02	γH, 2.32, 2.33; NH, 7.02, 7.69
Ile^3	8.68	4.16	1.83	γCH$_2$, 1.19, 1.51; γCH$_3$, 0.87; δH, 0.87
Ile^4	8.45	4.16	1.82	γCH$_2$, 1.16, 1.43; γCH$_3$, 0.84; δH, 0.84
Asn^5	8.68	4.66	2.61, 2.59	NH, 6.96, 7.64
Met^6	8.54	4.37	1.85, 1.95	γH, 2.36, 2.38; ϵH, nd
Trp^7	8.18	4.64	3.30	NH, 10.22; 2H, 7.25; 4H, 7.55; 5H, 7.11; 6H, 7.20; 7H, 7.45
Gln^8	8.00	4.18	1.83, 2.00	γH, 2.06, 2.17; NH, 6.93, 7.56
Glu^9	8.29	4.30	1.95, 2.06	γH, 2.35, 2.39
Val^{10}	8.33	4.07	2.10	γH, 0.94, 0.97
Gly^{11}	8.63	3.94		
Lys^{12}	8.21	4.27	1.74, 1.84	γH, 1.40; δH, 1.65; ϵH, 2.96; NH, 7.61
Ala^{13}	8.42	4.24	1.33	
Met^{14}	8.33	4.37	1.90	γH, 2.35, 2.43; ϵH, nd
Tyr^{15}	8.20	4.57	2.93, 3.06	2,6H, 7.05, 7.12; 3,5H, 6.78, 6.82
Ala^{16}	8.29	4.30	1.36	
Cys^{17}	8.38	4.55	2.94	
Thr^{18}	8.38	4.33	4.21	γH, 1.20
Arg^{19}	8.46	4.63	1.75, 1.82	γH, 1.64, 1.67; δH, 3.19; NH, 7.27
Pro^{20}		4.40	1.87, 2.28	γH, 2.00, 2.03; δH, 3.58, 3.74
Asn^{21}	8.71	4.67	2.78, 2.84	NH, 7.03, 7.73
Asn^{22}	8.60	4.72	2.77, 2.84	NH, 7.03, 7.73
Asn^{23}	8.61	4.76	2.80, 2.86	NH, 7.06, 7.70
Thr^{24}	8.23	4.30	4.23	γH, 1.21
Arg^{25}	8.42	4.34	1.77, 1.84	γH, 1.58; δH, 3.17; NH, 7.23
Lys^{26}	8.49	4.34	1.75, 1.83	γH, 1.46; δH, 1.68; ϵH, 2.99; NH, 7.60
Ser^{27}	8.52	4.48	3.84	
Ile^{28}	8.48	4.30	1.90	γCH$_2$, 1.19, 1.46; γCH$_3$, 0.91; δH, 0.86
Thr^{29}	8.42	4.35	4.16	γH, 1.19
Lys^{30}	8.57	4.40	1.76, 1.86	γH, 1.45; δH, 1.68; ϵH, 2.98; NH, 7.61
Gly^{31}	8.46	4.06, 4.15		
Pro^{32}		4.45	1.88, 2.28	γH, 1.97, 2.02; δH, 3.62, 3.65
Gly^{33}	8.64	3.91, 3.95		
Arg^{34}	8.25	4.33	1.74, 1.78	γH, 1.54, 1.56; δH, 3.16; NH, 7.23
Val^{35}	8.37	3.98	1.89	γH, 0.66, 0.89
Ile^{36}	8.37	4.11	1.75	γCH$_2$, 1.10, 1.36; γCH$_3$, 0.80; δH, 0.80
Tyr^{37}	8.47	4.63	2.85, 3.08	2,6H, 7.05, 7.12; 3,5H, 6.78
Ala^{38}	8.53	4.41	1.41	
Thr^{39}	8.27	4.36	4.30	γH, 1.23
Gly^{40}	8.16	3.81		

^{a} nd, not determined.

A second region in the SP10RF(A) segment may also exhibit a turn. In the sequence from position 19 to position 22, a *d*αN(*i*,*i*+2) NOE is observed between Pro^{20} and Asn^{22}, and the *d*Nδ(*i*,*i*+1) and *d*δN(*i*,*i*+1) NOEs between Pro^{20} and adjacent residues are weak. This is consistent with a β turn in which residues Arg^{19}–Asn^{22} comprise turn residues 1–4, respectively. The sequence Pro-Asn is frequently found in β turns in peptides (Dyson et al., 1988c).

A third conformational feature detected in the SP10RF(A) segment is at positions Ser^{27}-Ile^{28}-Thr^{29}. Three NOEs are observed between Ser^{27} and Thr^{29} which are due to *d*βN(*i*,*i*+2) (Figure 5), *d*βα(*i*,*i*+2), and *d*αN(*i*,*i*+2) connectivities. These NOEs suggest an unusual proximity of residues 27, 28, and 29, possibly arising from hydrogen bonds forming between the Ser and Thr side chains or between the side chains and backbone amide groups. A *d*βN(*i*,*i*+2) NOE has previously been ascribed to a kinked backbone conformation (Osterhout et al., 1989).

The backbone configuration from Ile^{36} to Ala^{38} has strong *d*αN(*i*,*i*+1) NOEs, suggesting an extended β conformation in this stretch. No medium- or long-range backbone NOEs are detected among these residues. However, the side chains of Val^{35} and Tyr^{37} are clustered. NOEs are seen between the upfield (0.66 ppm) methyl group of Val^{35} and all ring hydrogens of Tyr^{37} (Figure 7). The low chemical shift value of this methyl resonance compared to the other methyl group (0.89 ppm) is consistent with a ring shift due to proximity to an aromatic group. Also, the Tyr^{37} ring hydrogens (2,6 and 3,5) show an NOE with the Val^{35} β-hydrogen (not shown).

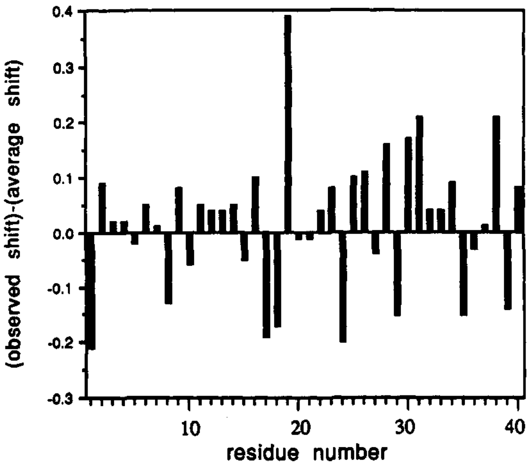

FIGURE 3: Plot of chemical-shift differences of C$^\alpha$ hydrogens in T1SP10RF(A). The ordinate is the difference of the observed shift minus the average shift for that amino acid, according to Wishart et al., (1991).

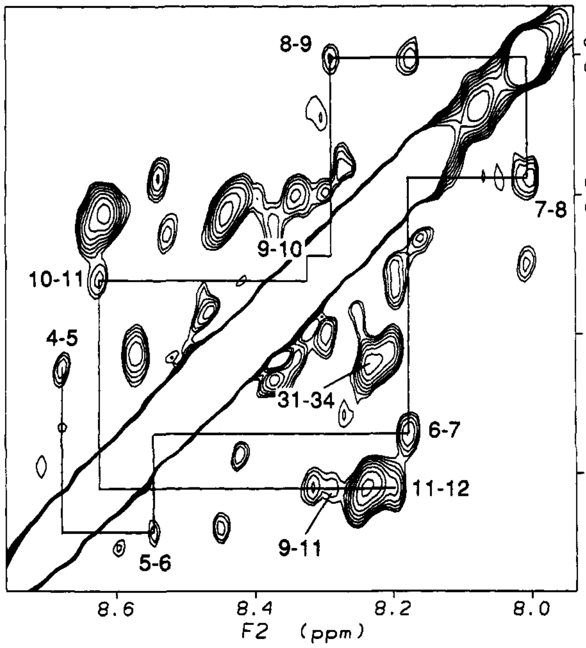

FIGURE 4: Partial trace of sequential backbone amide NOEs (indicated by residue number) in the T1 segment. Also indicated is a *d*NN(*i*,*i*+2) NOE (Glu^9–Gly^{11}) and a *d*NN(*i*,*i*+3) NOE (Gly^{31}–Arg^{34}).

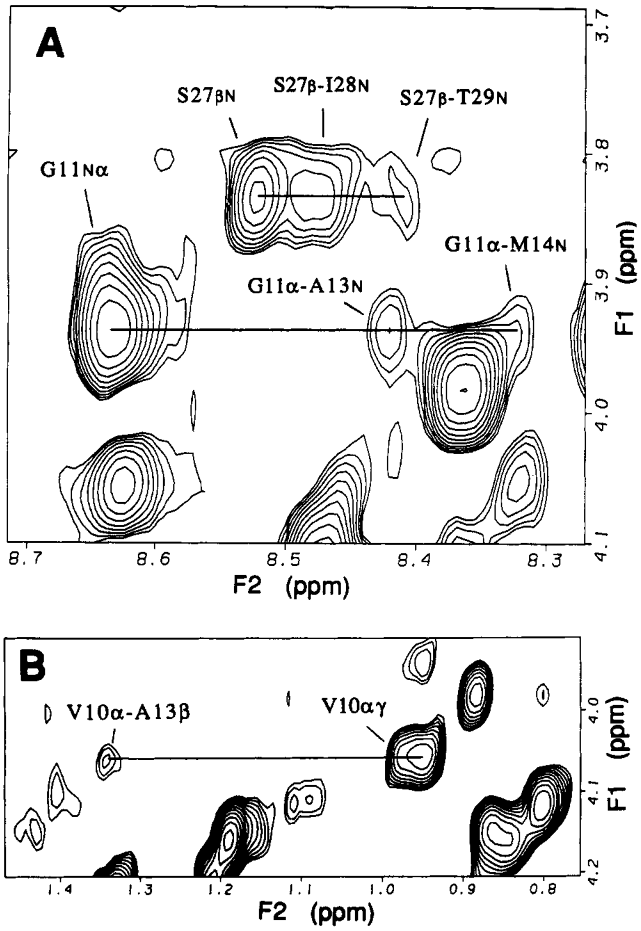

FIGURE 5: NOEs indicative of a full helical turn in the T1 segment. (A) $d\alpha N(i,i+3)$ between Gly^{11} and Met^{14}. Also shown is a $d\beta N(i,i+2)$ NOE between Ser^{27} and Thr^{29}. (B) $d\alpha\beta(i,i+3)$ NOE between Val^{10} and Ala^{13}.

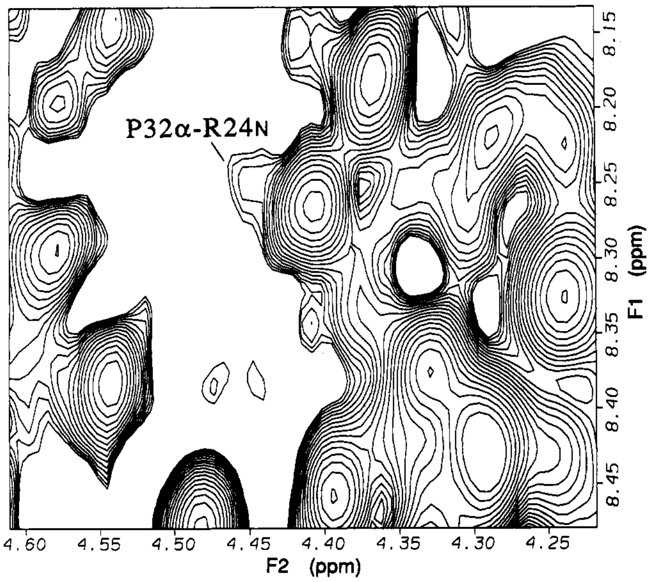

FIGURE 6: $d\alpha N(i,i+2)$ NOE between Pro^{32} and Arg^{34}.

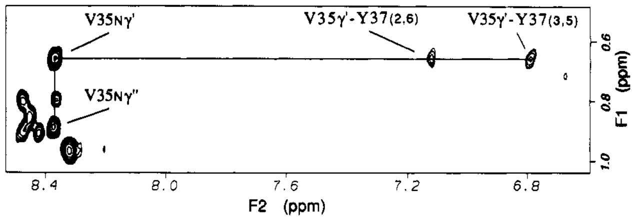

FIGURE 7: NOEs between the ring hydrogens of Tyr^{37} and the upfield methyl group of Val^{35}.

No NOEs are observed between either the Val^{35} or the Tyr^{37} side chains and that of Ile^{36}. These findings are expected of a β strand, where side chains protrude alternately along opposite sides of the backbone. A similar $(i,i+2)$ clustering of side chains has been observed for a peptide fragment of plastocyanin, and resembles the interaction found in native protein (Dyson et al., 1992b). The final three residues of the SP10RF(A) segment may adopt nascent helical conformations, as suggested by the $d\alpha N(i,i+2)$ NOE between Ala^{38} and Gly^{40} (Figure 2).

Disulfide Cross-Linked Peptide. T1SP10RF(A) peptide dissolved in buffer lacking DTT was found by electrospray mass spectrometry to consist largely (87% by weight) of a 8992-Da species, i.e., 2 mass units less than twice the monomeric mass. When this solution was made 10 mM in DTT, only the species of monomeric mass (4497 Da) was detected. Thus, the peptide appeared to form dimers by disulfide bond formation through cysteine residues in the absence of DTT. The molar ratio of monomer to dimer was approximately 1:3.3. NMR spectra of the mixed monomer/dimer solution were nearly identical to those of peptide in 10 mM DTT, wherein no dimer was present. The chemical shifts of all hydrogens, including the βCH_2 of Cys^{17}, were identical or shifted upfield by less than 0.02 ppm in the absence of DTT. Three exceptions were noted: the backbone amide hydrogens of Glu^9, Val^{10}, and Gly^{11} were shifted upfield by 0.03 ppm. The shift of these three hydrogens may be due to a lower pH (3.8 vs 4.2) in the solution lacking DTT, which would cause differential protonation of the Glu^9 side-chain carboxyl group. NOESY spectra of dimerized and monomeric peptide were almost identical, with some variation in the intensity of certain NOEs. A few NOEs were unique or much more intense in the solution of partially dimerized RF (Figure 8). Most of these occur between residues in the vicinity of Cys^{17} and thus probably arise from disulfide-linked peptides. An example of two such NOEs, between Ala^{13} βCH_3 and the ring hydrogens of Tyr^{15}, is shown in Figure 8B. The interpretation of interchain versus intrachain NOEs in Figure 8A is based on the assumption that a long-range intrachain NOE is less likely than a potentially short-range interchain NOE. For example, the NOEs due to Tyr^{15}ringH–Pro$^{20}\alpha$H and Tyr^{15}ringH–Arg$^{19}\alpha$H would be $(i,i+5)$ and $(i,i+4)$, respectively, if assigned to intrachain interaction, distances which are not commonly detected in monomeric linear peptide fragments. Given the uncertainty of assigning NOEs to interchain or intrachain interactions, it is presumptive to propose particular conformations which would account for them. Another dimer-specific NOE is due to Ser^{27}NH–Thr$^{29}\gamma CH_3$ (not shown). It is not obvious how disulfide cross-linking or a slight difference in pH could account for this NOE. Two other NOEs which are greatly diminished in the DTT solution occur at 1.34/4.41 ppm (possibly between Ala$^{13}\beta CH_3$ or Ala$^{16}\beta CH_3$ and Pro$^{20}\alpha$H) and 1.20/4.41 ppm (between Thr$^{18}\gamma CH_3$ and Pro$^{20}\alpha$H or Thr$^{29}\gamma CH_3$ and Lys$^{30}\alpha$H). Hence, these two NOEs might also be due to conformations induced by cross-linking, and possibly arise from interchain interactions.

cis Peptide Bonds. A number of $d\alpha\alpha(i,i+1)$ NOEs are observed which are likely due to *cis* peptide bonds. Such NOEs occur in the following pairs of residues: Ile^4-Asn^5, Val^{10}-Gly^{11}, Gly^{11}-Lys^{12}, Pro^{32}-Gly^{33}, and Tyr^{37}-Ala^{38} (Figure 2). All these NOEs are weak, as expected, since in an unconstrained peptide fragment *trans* peptide bonds are

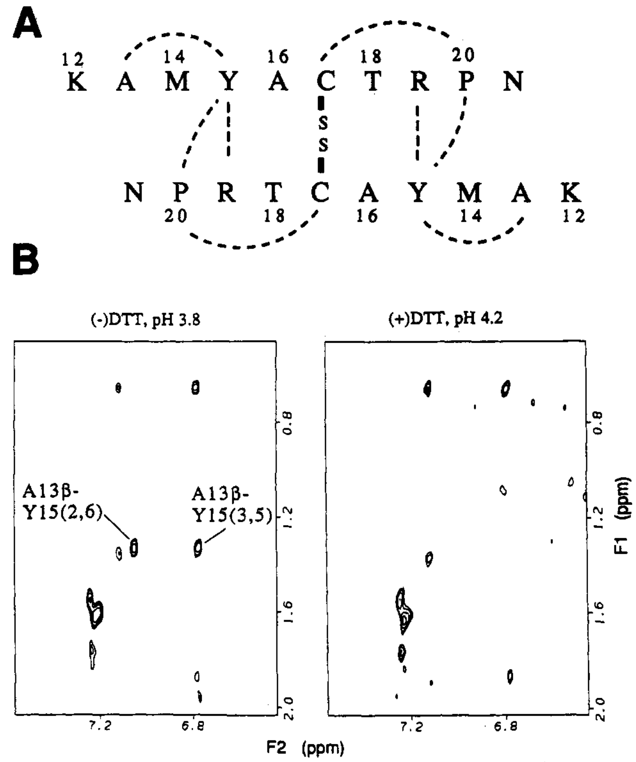

FIGURE 8: NOEs of T1SP10RF(A) that are affected by DTT. (A) The region surrounding Cys^{17} is shown as a disulfide cross-linked dimer. Dashed lines indicate interresidue NOEs which disappear or are much less intense upon reduction of the disulfide bond. The particular NOEs are Cys$^{17}\alpha$–Pro$^{20}\delta$, Ala$^{13}\beta$–Tyr^{15}(2,6), Ala$^{13}\beta$–Tyr^{15}(3,5), Ala$^{13}\alpha$–Tyr^{15}(2,6), Tyr^{15}(2,6)–Pro$^{20}\alpha$, Tyr^{15}(3,5)–Pro$^{20}\alpha$, Tyr^{15}(2,6)–Arg$^{19}\alpha$, and Tyr^{15}(3,5)–Arg$^{19}\alpha$. (B) Illustration of two NOE cross-peaks which depend on the presence or absence of DTT. Both spectra were processed identically and are shown at the same contour level.

energetically favored. For peptides containing proline, the X–Pro bond exists substantially in the *cis* isomer, but $d\alpha\alpha(i,i+1)$ NOEs due to Arg^{19}-Pro^{20} and Gly^{31}-Pro^{32} are not detected. The apparent adoption of tight turns by these residues may reduce the fraction of time they exist in the *cis* configuration.

DISCUSSION

Using NMR spectroscopy, we have detected preferred conformations in T1SP10RF(A) in aqueous solution. To draw inferences on the immunogenicity of a peptide, it is important to compare the observed conformations with the structure of the peptide in the native protein. The three-dimensional structure of gp120 is yet undetermined, so comparison of the native versus the solution structure of its peptide segments is not possible. Nevertheless, predictions have been made of secondary structure in the V3 loop based on sequence analysis of many isolates (LaRosa et al., 1990). A probable feature in this region is a β turn (type II) at the Gly-Pro-Gly-Arg sequence, preceded and followed by approximately five residues of β strand. Our results provide evidence that a turn (type I) followed by a stretch of β stand is present in the RF peptide in solution. NMR studies of V3 peptides from the MN and III$_B$ isolates, the sequences of which differ considerably from RF in this region (RF, SITKGPGRVIYA; MN, RIHIGPGRAFYT; III$_B$, RIRQGPGRAFYT), also showed that a type I turn is present at the Gly-Pro-Gly-Arg sequence and is flanked by extended conformation (Chandrasekhar et al., 1991; Zvi et al., 1992). Although the backbone conformations of these V3 peptides are similar in this region, sequence differences lead to the frequent elicitation of strain-specific antibodies (Palker et al., 1988; Rusche et al., 1988). However, neutralizing antibodies to the V3 loop that display lower strain-specificity have been raised with a hybrid peptide analogous to T1SP10RF(A) but bearing the V3 sequence of strain MN (Haynes et al., 1993). These antibodies neutralized HIV strains III$_B$ and RF in addition to MN, but not strain Du4489-5, which differs from MN by one replacement (Du4489-5: RIPIGPGRAFYT). This implies that the dominant epitope in the immunogenic MN peptide is a conserved conformational motif, presumably the turn formed by Gly-Pro-Gly-Arg, which is of different structure in the native V3 loop of strain Du4489-5.

The T1 segment in native gp120 elicits neutralizing antibodies in mice (Lasky et al., 1987; Sun et al., 1989). However, in contrast to V3 peptides, free T1 peptide is not immunogenic for neutralizing antibodies (Ho et al., 1987; Palker et al., 1989b; Haynes et al., 1993). Several reasons could account for this finding, but we raise the possibility that it is due to a preference for highly nonnative conformations by the T1 region in solution, as either free or hybrid peptides. This hypothesis derives from the finding that a 39-residue stretch of gp120 which includes T1 has significant homology to the first constant domain (C$_H$1) of the α2 heavy chain of human IgA (Maddon et al., 1986). By analogy to the structure of a mouse IgA (Segal et al., 1974), the gp120-homologous portion of IgA α2 corresponds to the final three β strands of C$_H$1 and their short connecting loops (Beale & Feinstein, 1976). The T1 segment of gp120 would correspond largely to the putative final (seventh) β strand of the IgA C$_H$1 domain. Thus, T1 may form a β strand in gp120, in contrast to its propensity to exhibit helical conformations as a free peptide. If this is the case, then the set of antibodies elicited by the peptide may not include those which bind to its native conformation.

In addition to B-cell epitopes, peptide T1SP10RF(A) also contains epitopes recognized by T lymphocytes. Three segments of RF, one in T1 and two in SP10RF(A), each contain a separately defined T-cell antigen (Figure 1). The factors determining which fragments of a protein become associated with MHC molecules for presentation to T cells are not fully defined. There is a relation between particular MHC alleles and the sequences of peptides which they bind (Falk et al., 1991). The peptides are generally nonamers in which certain positions are dominated by particular residues. It is not known if this is the sole determinant of antigen presentation as there are numerous steps in antigen processing, any of which might impart selectivity for or against particular fragments. One selective factor may be the secondary structure of the peptide fragment. Various studies have implicated features such as amphipathic helix (Spouge et al., 1987; Margalit et al., 1987; Rothbard et al., 1988), α helix, or β strand (Schrier et al., 1989), lack of a tight turn (Schrier et al., 1989), and lack of random-coil conformation (Spouge et al., 1987) in selection for particular fragments. Hence, we examined the conformations of T-cell antigenic portions of peptide T1SP10RF(A) in light of these suggestions.

The T1 region was originally predicted to be a T-cell antigen on the hypothesis that such antigens fold as amphipathic helices (Cease et al., 1987). Consistent with this, we found that much of T1 adopted conformations typical of helices, and in one portion exhibited a tendency to form a full helical turn. A T-helper epitope of myoglobin also has a propensity to adopt helical conformations as a peptide (Lark et al., 1989; Waltho et al., 1989). On the other hand, there are T-cell antigenic peptides which, according to circular dichroism studies, do not adopt helical conformations (Abergel et al., 1989). Similarly, we found that the segment of T1SP10RF(A) from Cys^{17} to Ile^{28} has little helical propensity, yet contains a T-cell antigen. Furthermore, when modeled as a helix, this region is not amphipathic (not shown). The segment from Ser^{27} to Gly^{40} may exhibit helical propensity only in the last three residues. Much of this segment exhibits other conformations, such as the kink from Ser^{27} to Thr^{29}, a tight turn from Gly^{31} to Arg^{34}, and a β strand from Val^{35} to Ala^{38}. The likely presence of a turn in the Ser^{27}–Gly^{40} segment and the possibility of one in the Cys^{17}–Ile^{28} segment contrast with the proposal of a bias against tight turns in T-cell antigenic peptides (Schrier et al., 1989). A general feature common to all three antigenic segments is simply the presence of medium-range interresidue interactions, indicating a lack of random-coil conformation. A selection against random-coil conformations in T-cell antigens has been proposed (Spouge et al., 1987). Perhaps peptides with highly random conformations are more susceptible to complete proteolytic degradation.

The ability to detect conformations in peptide T1SP10RF(A) provides a basis for studies on similar peptides derived from gp120 sequences of other strains of HIV. Given that hybrid peptides are more immunogenic than separate epitopes, it will be of interest to determine whether particular conformations are affected by covalent linkage to other segments. As comparative data on immune responses to such peptides become available, it should be possible to draw more detailed inferences on the role of solution conformation in peptide immunogenicity.

ACKNOWLEDGMENT

We thank Ronald Venters and David Hyre for assistance with obtaining and processing NMR spectra, David Myers for discussion of results on a related hybrid peptide, Craig Heinly for peptide purification, and Robert Stevens and David Millington for performing mass spectrometric analyses. Also acknowledged is the Duke NMR Center, whose instrumentation was purchased with grants from the NIH, NSF, and the North Carolina Biotechnology Center.

REFERENCES

Abergel, C., Loret, E., & Claverie, J.-M. (1989) *Eur. J. Immunol. 19*, 1969–1972.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Beale, D., & Feinstein A. (1976) *Q. Rev. Biophys. 9*, 135–180.

Cease, K. B., Margalit, J., 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. U.S.A. 84*, 4249–4253.

Chandrasekhar, K., Profy, A. T., & Dyson, H. J. (1991) *Biochemistry 30*, 9187–9194.

Dadaglio, G., Leroux, A., Langlade-Demoyen, P., Bahraoui, E. M., Traincard, F., Fisher, R., & Plata, F. (1991) *J. Immunol. 147*, 2302–2309.

DeLisi, C., & Berzofsky, J. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7048–7052.

Dyson, H. J., & Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519–538.

Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature 318*, 480–483.

Dyson, H. J., Lerner, R. A., & Wright, P. E. (1988a) *Annu. Rev. Biophys. Biophys. Chem. 17*, 305–374.

Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.

Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988c) *J. Mol. Biol. 201*, 161–200.

Dyson, H. J., Satterthwait, A. C., Lerner, R. A., & Wright, P. E. (1990) *Biochemistry 29*, 7828–7837.

Dyson, H. J., Norrby, E., Hoey, K., Parks, D. E., Lerner, R. A., & Wright, P. E. (1992a) *Biochemistry 31*, 1458–1463.

Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H.-C., Lerner, R. A., & Wright, P. E. (1992b) *J. Mol. Biol. 226*, 819–835.

Falk, K., Rötzschke, O., Stevanovic, S., Jung, G., & Rammensee, H.-G. (1991) *Nature 351*, 290–296.

Germain, R. N., & Margulies, D. H. (1993) *Annu. Rev. Immunol. 11*, 403–450.

Goudsmit, J., Debouck, C., Meloen, R. H., Smit, L., Bakker, M., Asher, D. M., Wolff, A. V., Gibbs, C. J., Jr., & Gajdusek, D. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4478–4482.

Gras-Masse, H., Jolivet, M., Drobecq, H., Aubert, J. P., Beachey, E. H., Audibert, F., Chedid, L., & Tartar, A. (1988) *Mol. Immunol. 25*, 673–678.

Hart, M. K., Palker, T. J., Matthews, T. J., Langlois, A. J., Lerche, N. W., Martin, M. E., Scearce, R. M., McDanal, C., Bolognesi, D. P., & Haynes, B. F. (1990) *J. Immunol. 145*, 2677–2685.

Hart, M. K., Weinhold, K. J., Scearce, R. M., Washburn, E. M., Clark, C. A., Palker, T. J., & Haynes, B. F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9448–9452.

Haynes, B. F., Torres, J. V., Langlois, A. J., Bolognesi, D. P., Gardner, M. B., Palker, T. J., Scearce, R. M., Jones, D. M., Moody, M. A., McDanal, C., & Matthews, T. J. (1993) *J. Immunol. 151*, 1646–1653.

Ho, D. D., Sarngadharan, M. G., Hirsch, M. S., Schooley, R. T., Rota, T. R., Kennedy, R. C., Chanh, T. C., & Sato, V. L. (1987) *J. Virol. 61*, 2024–2028.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Phys. Chem. 71*, 4546–4553.

Lark, L. R., Berzofsky, J. A., & Gierasch, L. M. (1989) *Peptide Res. 2*, 314–321.

LaRosa, G. J., Davide, J. P., Weinhold, K., Waterbury, J. A., Profy, A. T., Lewis, J. A., Langlois, A. J., Dreesman, G. R., Boswell, R. N., Shadduck, P., Holley, L. H., Karplus, M., Bolognesi, D. P., Matthews, T. J., Emini, E. A., & Putney, S. (1990) *Science 249*, 932–935.

Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory T., & Capon, D. J. (1987) *Cell 50*, 975–985.

Lerner, R. A. (1984) *Adv. Immunol. 36*, 1–44.

Levitt, M. H., Freeman, R., & Frenkiel, T. (1982) *J. Magn. Reson. 47*, 328–330.

Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R., & Axel, R. (1986) *Cell 47*, 333–348.

Margalit, H., Spouge, J. L., Cornette, J. L., Cease, K. B., Delisi, C., & Berzofsky, J. A. (1987) *J. Immunol. 138*, 2213–2229.

Ni, F., Ripoll, D. R., & Purisima, E. O. (1992) *Biochemistry 31*, 2545–2554.

Osterhout, J. J., Baldwin, R. L., York, E. J., Stewart, J. M., Dyson, H. J., & Wright, P. E. (1989) *Biochemistry 28*, 7059–7064.

Palker, T. J., Clark, M. E., Langlois, A. J., Matthews, T. J., Weinhold, K. J., Randall, R. R., Bolognesi, D. P., & Haynes, B. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1932–1936.

Palker, T. J., Matthews, T. J., Langlois, A., Kim, J. E., Berzofsky, J. A., Bolognesi, D. P., & Haynes, B. F. (1989a) *UCLA Symp. Mol. Cell. Biol., New Ser. 119*.

Palker, T. J., Matthews, T. J., Langlois, A., Tanner, M. E., Martin, M. E., Scearce, R. M., Kim, J. E., Berzofsky, J. A., Bolognesi, D. P., & Haynes, B. F. (1989b) *J. Immunol. 142*, 3612–3619.

Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., & Wong-Stahl, F. (1985) *Nature 313*, 277–284.

Rothbard, J., & Taylor, W. (1988) *EMBO J. 7*, 93–100.

Rothbard, J. B., Lechler, R. I., Howland, K., Bal, V., Eckels, D. D., Sekaly, R., Long, E. O., Taylor, W. R., & Lamb, J. R. (1988) *Cell 52*, 515–523.

Rusche, J. R., Javaherian, K., McDanal, C., Petro, J., Lynn, D. L., Grimaila, R., Langlois, A., Gallo, R. C., Arthur, L. O., Fischinger, P. J., Bolognesi, D. P., Putney, S. D., & Matthews, T. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3198–3202.

Schrier, R. D., Gnann, J. W., Jr., Landes, R., Lockshin, C., Richman, D., McCutchan, A., Kennedy, C., Oldstone, M. B. A., & Nelson, J. A. (1989) *J. Immunol. 142*, 1166–1176.

Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M., & Davies, D. R. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4298–4302.

Shirai, M., Pendelton, C. D., Ahlers, J., Takeshita, T., Newman, M., & Berzofsky, J. A. (1994) *J. Immunol. 152*, 549–556.

Spouge, J. L., Guy, H. R., Cornett, J. L., Margalit, H., Cease, K., Berzofsky, J. A., & DeLisi, C. (1987) *J. Immunol. 138*, 204–212.

Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolff, H., Parks, E. S., Josephs, S. F., Gallo, R. C., & Wong-Staal, F. (1986) *Cell 45*, 637–648.

Sun, N.-C., Ho, D. D., Sun, C. R. Y., Liou, R.-S., Gordon, W., Fung, M. S.-C., Li, X.-L., Ting, R. C., Lee, T.-H., Chang, N. T., & Chang, T.-W. (1989) *J. Virol. 63*, 3579–3585.

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. U.S.A. 85*, 3105–3109.

Takahashi, H., Houghten, R., Putney, S. D., Margulies, D. H., Moss, B., Germain, R. N., & Berzofsky, J. A. (1989) *J. Exp. Med. 170*, 2023–2035.

Takahashi, H., Nakagawa, Y., Pendleton, C., Houghten, R. A., Yokomuro, K., Germain, R. N., & Berzofsky, J. A. (1992) *Science 255*, 333–336.

Tsang, P., Rance, M., Fieser, T. M., Ostresh, J. M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1992) *Biochemistry 31*, 3862–3871.

Wahren, B., Rosen, J., Sandström, E., Mathiesen, T., Modrow, S., & Wigzell, H. (1989) *J. AIDS 4*, 448–456.

Waltho, J. P., Feher, V. A., Lerner, R. A., & Wright, P. E. (1989) *FEBS Lett. 250*, 400–404.

Williamson, M. P., Hall, M. J., & Handa, B. K. (1986) *Eur. J. Biochem. 158*, 527–536.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.

Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.

Zvi, A., Hiller, R., & Anglister, J. (1992) *Biochemistry 31*, 6972–6979.