The Membrane-Proximal Tryptophan-Rich Region of the HIV Glycoprotein, gp41, Forms a Well-Defined Helix in Dodecylphosphocholine Micelles[†,‡]

David J. Schibli,[§] Ronald C. Montelaro,[||] and Hans J. Vogel*[,§]

Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, Canada, and Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261

Received March 29, 2001; Revised Manuscript Received June 11, 2001

ABSTRACT: The membrane-proximal tryptophan-rich region of the HIV transmembrane glycoprotein, gp41, plays an important role in the membrane fusion reaction. Using NMR spectroscopy, we have studied the tertiary structure of a synthetic 19-residue amidated peptide (NH_2-KWASLWNWFNITNWLWYIK-$CONH_2$) corresponding to this region in membrane-mimetic environments. Initial experiments in sodium dodecyl sulfate/H_2O micelles and trifluoroethanol gave poor results, because of low solubility. However, in dodecylphosphocholine micelles, we obtained excellent 500 and 800 MHz NMR spectra, suggesting that the peptide has a preference for a zwitterionic membrane-like environment. The final NMR structures demonstrated a well-defined helical peptide with a backbone rmsd of 0.47 ± 0.18 Å. Four of the five tryptophan residues, as well as the tyrosine residue, formed a "collar" of aromatic residues along the axial length of the helix. By analogy to related tryptophan-rich antimicrobial peptides, the structure indicates that the aromatic residues of the HIV peptide are positioned within the membrane−water interface of a phospholipid bilayer. This is confirmed by the observation of direct NOEs between the aromatic residues of the peptide to the headgroup and interfacial protons of prototonated dodecylphosphocholine. The bulk of the polar residues are positioned on one face of this structure, with the hydrophobic phenylalanine side chain on the opposing face, forming an amphipathic structure. This work shows that the Trp-rich membrane-proximal region of HIV and related viruses can bind to the surfaces of zwitterioninc membranes in a "Velcro-like" manner.

Membrane-encapsulated viruses have a ubiquitous membrane fusion protein, and while the fusion mechanism may differ between viruses, in all cases these proteins possess a receptor binding subunit and a fusion subunit (reviewed in ref *1*). The fusion protein of HIV is initially expressed as the precursor protein gp160, which is subsequently proteolytically cleaved into two associated proteins, the receptor binding protein, gp120, and the fusion protein, gp41 (*2*). The gp41 subunit is a transmembrane protein, anchoring the gp120−gp41 complex to the membrane. When gp120 binds to the target cell CD4 receptor (*3*, *4*) and a chemokine coreceptor (*5*), gp41 is thought to undergo a conformational change (*6*−*8*; reviewed in ref *9*) mediating the fusion of the viral and target cell membranes.

It is expected that in the native gp41−gp120 complex, gp41 is present in a metastable conformation with the N-terminal fusion peptide buried near the postulated C-terminal transmembrane domain (*9*, *10*) similar to the influenza fusion peptide in the inactive hemaglutinin (HA)[1] structure (*11*). The gp120−CD4-induced conformational change leads to the "prehairpin intermediate" in which the fusion peptide is extended toward the target membrane and the N- and C-helices have become separated (*9*). The fusion-active core structure of the ectodomain of HIV gp41, thought to be responsible for the final fusion of the target and viral membranes, has been determined using X-ray crystallography, demonstrating a coiled-coil, α-helical hairpin trimer (*10*, *12*, *13*). The three N-helices (Figure 1) form a coiled-coil core with the three C-helices wrapped around the central N-helices. This coiled coil is thought to be analogous to the low-pH structure of influenza HA (*14*). The HIV gp41 ectodomain structures that have been determined to date lack the N-terminal fusion peptide, various connecting loops, and the transmembrane and cytoplasmic domains. The structure of the ectodomain of SIV gp41 containing the connecting

[†] This work was supported by an operating grant from the Canadian Institute for Health Research to H.J.V. and an NIH grant to R.C.M. (Grant 5RO1 AI36198). D.J.S. is supported by a NSERC studentship and an Alberta Heritage Foundation for Medical Research incentive award. H.J.V. holds a Scientist Award from the Alberta Heritage Foundation for Medical Research.

[‡] The atomic coordinates for the gp41W peptide in DPC micelles (1JAU, ensemble of the 40 lowest-energy structures; 1JAV, minimized average structure) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers, the State University of New Jersey, New Brunswick, NJ. The chemical shifts and table of assignments will be deposited in the BioMagResBank (University of Madison−Wisconsin, Madison, WI).

* To whom correspondence should be addressed. Telephone: (403) 220-6006. Fax: (403) 289-9311. E-mail: vogel@ucalgary.ca.

[§] University of Calgary.

[||] University of Pittsburgh School of Medicine.

[1] Abbreviations: CD, circular dichroism; DPC, dodecylphosphocholine; SDS, sodium dodecyl sulfate; TFE, trifluoroethanol; gp, glycoprotein; gp41W, HIV_{HXB2} gp41 (residues 665−684); NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; TOCSY, total correlation spectroscopy; LUV, large unilamellar vesicle; DOPE, dioleoylphosphatidylethanolamine; DOPC, dioleoylphosphatidylcholine; HA, hemagglutinin.

10.1021/bi010640u CCC: $20.00
Published on Web 07/18/2001

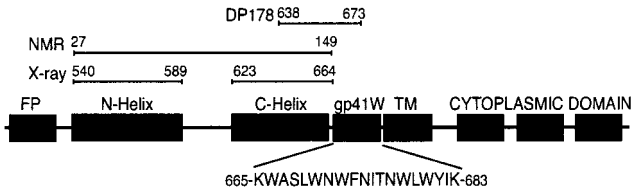

FIGURE 1: Schematic representation of the postulated secondary structure of gp41. The location and sequence of the Trp-rich membrane-proximal region used in this study are indicated. Additionally, the locations of other areas of interest are highlighted: the fusion inhibitory peptide, DP178 (*25*); the sequence of the SIV gp41 ectodomain construct studied by NMR spectroscopy (*15*); and the sequence of the longest HIV gp41 ectodomain construct studied by X-ray crystallography (*10*). The numbering scheme is based on the HIV gp160 sequence, except for the SIV sequence, which is based on SIV e-gp41 (*15*).

loops between the N- and C-terminal helices was determined using NMR spectroscopy, confirming the connectivity of the N- and C-terminal helices (*15*). This structure resembled those determined for the HIV ectodomain using X-ray crystallography, yet there is still no structural information available for the region connecting the ectodomain to the transmembrane and cytoplasmic domains (Figure 1).

The membrane-proximal region of the ectodomain of HIV-1_{HXB2} gp41 is rich in the amino acid tryptophan. This region has recently been reported to be essential for the membrane fusion mechanism of HIV (*16*). Specifically, the Trp residues have been demonstrated to be necessary for the formation of expanding fusion pores and syncytia formation (*17*). Salzwedel et al. (*16*) have recently suggested that the membrane-proximal region may play a role in the fusion of the target and viral membranes. It has been reported that the side chain of Trp is unique and that it has a preference for being inserted into lipid bilayers at the membrane−water interface (*18*−*22*). A synthetic peptide corresponding to the membrane-proximal region of gp41 (residues 664−683) was recently demonstrated to induce membrane leakage and lipid mixing of DOPC/DOPE/cholesterol LUVs (*23*, *24*), supporting the theory that this peptide may aid in disrupting the viral membrane during fusion. The critical role that this region plays in membrane fusion is emphasized by the sequence overlap of the gp41 DP178 inhibitory peptide (residues 638−672) with the N-terminus of the Trp-rich region (*25*). Additionally, the membrane fusion inhibitory antibody epitope, 2F5, is also located at the N-terminus of this region (*26*).

The mechanism by which gp41 induces the fusion of the target and viral membrane has been related to the pH-dependent endosomal fusion mechanism of influenza HA (*10*). While the similarities between gp41 and HA do hint at a related mechanism, the pH-independent mechanism of gp41 suggests that there may be some significant structural and mechanistic differences between the two proteins. In this study, we describe the NMR solution structure of a C-terminal amidated synthetic peptide corresponding to the membrane-proximal tryptophan-rich region of HIV_{HXB2} gp41 (gp41W; Figure 1) bound to a membrane-mimetic DPC micelle. Because of the large number of Trp residues, and an apparent micelle/peptide molecular mass approaching 20 kDa, spectral overlap was significant, necessitating the use of an 800 MHz NMR spectrometer in obtaining maximum resolution.

MATERIALS AND METHODS

Materials. The synthetic gp41W peptide (KWASLWNWFNITNWLWYIK-NH_2) used in the study was produced by the synthetic peptide core of the Biomedical Research Support Facility of the University of Pittsburgh (Pittsburgh, PA). Peptide production procedures have been described in detail previously (*27*). Briefly, the gp41W peptide was synthesized as a C-terminal amide using standard Fmoc synthesis protocols and purified to homogeneity by reverse phase high-performance liquid chromatography. The predicted peptide sequence was confirmed by mass spectrometry, and peptide concentrations were quantified using a sensitive ninhydrin assay. Perdeuterated DPC and D_2O were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Protonated DPC was purchased from Avanti Polar Lipids (Alabaster, AL).

CD Spectroscopy. CD experiments were performed on a Jasco J-715 CD spectrophotometer with a 1 mm path length cylindrical cuvette. Samples were measured between 185 and 255 nm, with a 0.2 nm step resolution, a measurement speed of 50 nm/min, a 2 s response time, and a 1 nm bandwidth. gp41W samples in DPC were prepared from a stock solution of gp41W in 100 mM DPC and diluted to 40 μM gp41W in 50 mM DPC and 10 mM Tris (pH 7.0). Ten spectra were collected and averaged at room temperature. Before smoothing was carried out, a blank, lacking peptide, was subtracted.

NMR Spectroscopy. An NMR sample containing gp41W in DPC was made by dissolving approximately 4.4 mg of gp41W in a 9:1 H_2O/D_2O mixture. The peptide was solubilized upon addition of 200 mM DPC-d_{38}, and the pH was adjusted to 3.51. The concentration was estimated to be 1.6 mM using UV absorption at 280 nm and a calculated molar extinction coefficient of 29 730 M^{-1} cm^{-1} based on the number of Trp and Tyr residues. Any potential effects of DPC-d_{38} on the UV absorption were not accounted for.

Initial one-dimensional ^1H NMR spectra were acquired on a Bruker Avance 500 MHz NMR spectrometer at 18, 25, 37, and 42 °C. The 37 °C spectra gave the optimum balance between line width and chemical shift distribution, so all further spectra were acquired at this temperature. Water suppression for the one-dimensional ^1H spectra was achieved using low-power presaturation during the relaxation delay. DQF-COSY, TOCSY (with a 40 ms mixing time), and NOESY spectra (with mixing times of 100 and 50 ms) were acquired on a Bruker Avance 500 MHz NMR spectrometer. Water suppression was achieved in the TOCSY and NOESY spectra using excitation sculpting (*28*), and in the COSY spectra using the 3−9−19 watergate pulse sequence (*29*) with gradients and presaturation during the relaxation delay. In all of the spectra, the number of data points was 2048 and 512 in the F_2 and F_1 dimensions, respectively, with sweep widths of 7046 Hz. Spectra were zero-filled to a 2K $\times$ 2K matrix and multiplied by a shifted sine-bell curve. NOESY and TOCSY spectra were also obtained on a Varian 800 MHz NMR spectrometer at NANUC, the National High Field Nuclear Magnetic Resonance Centre (University of Alberta, Edmonton, AB). For the 800 MHz data, water suppression was achieved using a watergate pulse sequence for both the NOESY and TOCSY spectra (ProteinPack, Varian Inc., Palo Alto, CA). The number of data points was 2048 and 1706 for the 800 MHz NOESY spectra and 4096 and 2048 for the TOCSY spectra for the F_2 and F_1

dimensions, respectively. The sweep widths of both spectra were 12 000 and 10 000 Hz in the F_2 and F_1 dimensions, respectively. Spectra were processed in a manner similar to that discussed above. All data were processed on a Silicon Graphics Indy workstation using NMRPipe (*30*).

To study the backbone amide exchange rates, the gp41W, DPC NMR sample was lyophilized and redissolved in 99.9% D_2O and a series of one-dimensional ^{1}H NMR spectra were acquired as soon as possible and over the subsequent 24 h to observe slowly exchanging amides.

To determine the sites of interaction between the peptide side chains and DPC, 2 mM gp41W peptide in a 9:1 H_2O/D_2O mixture was solubilized in 50 mM protonated DPC ([^{1}H]DPC) micelles. Two-dimensional ^{1}H NOESY spectra were acquired at 37 °C, with mixing times of 500, 300, and 100 ms on a Bruker Avance 500 MHz NMR spectrometer equipped with a Bruker Cryo-probe. Two-dimensional ^{1}H TOCSY spectra with mixing times of 60 ms were obtained for the gp41W/[^{1}H]DPC sample as well as for 20 mM [^{1}H]-DPC. Spectra were acquired and processed as described above, except that only 16 scans were acquired per experiment. Additionally, two-dimensional natural abundance ^{13}C−^{1}H HSQC and ^{13}C−^{1}H HSQC-TOCSY spectra were acquired for 20 mM [^{1}H]DPC with 1024 × 256 data points and 16 scans.

Structure Calculation. NMR spectra were analyzed with the program NMRView 4.1.3 (*31*). The gp41W chemical shifts were assigned using standard techniques (*32*). Distance restraints used in the structure calculations were based on NOE peak intensities, and were classified as strong, medium, or weak restraints, corresponding to distance restraints of 1.8−2.8, 1.8−3.4, and 1.8−5.0 Å, respectively. The distance restraints were calibrated using various intraresidue NOEs in the five tryptophan rings. Where possible, atoms were stereospecifically assigned on the basis of NOE patterns. Pseudoatom corrections were applied to degenerate and non-stereospecifically assigned protons (*32*). The ϕ dihedral angle was constrained between −35° and −180° (usually the only range considered in NMR-derived structures).

The initial peptide structures were determined from an extended peptide structure using molecular dynamics with the program CNS (*33*). The starting structure went through a high-temperature annealing stage and two slow-cooling annealing stages, and then was subjected to 200 steps of Powell energy minimization. The peptide structures were further refined using the program ARIA (*34*−*37*). ARIA enabled the incorporation of ambiguous distance restraints and calibration of the NOE restraints using automated matrix analysis as implemented by the program. Aria runs were performed using the default parameters with eight iterations. Twenty structures were generated each round, and the 10 lowest-energy structures were carried on to the next iteration. In the final iteration, 200 structures were generated, and the 40 lowest-energy structures were retained as the final structures.

RESULTS

CD Spectroscopy. The CD spectra of gp41W in DPC micelles has a positive peak at approximately 195 nm, with major negative peaks at 208 and approximately 220 nm (Figure 2). The spectrum resembles that of a typical α-helical peptide, with a positive maximum at 194 nm and negative

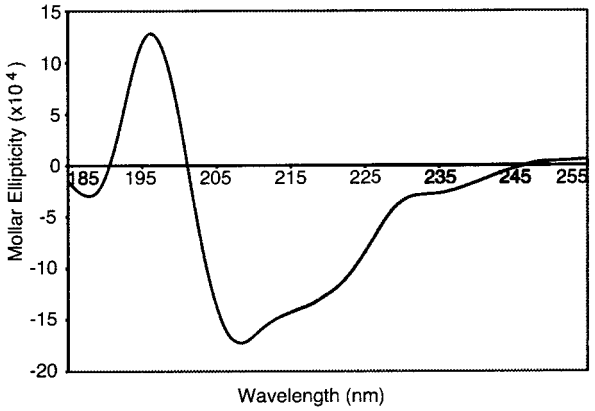

FIGURE 2: CD spectrum of 40 μM gp41W in 50 mM DPC micelles and 10 mM Tris (pH 7.0) at 25 °C.

maxima at approximately 210 and 220 nm. It should be noted that in a pure α-helical peptide the negative minima at 208 and 222 nm are of approximately equal intensity, but as can be seen in Figure 2, the intensity of the 208 nm peak is greater than that of the 220 nm peak. This difference could be due to multiple factors, such as the presence of random-coil or turn structures in the peptide. Additionally, the presence of aromatic residues, specifically tryptophan, in a peptide complicates the interpretation of CD spectra, with either positive or negative contributions near 220 nm, depending on the orientation of the indole ring (*38*). The secondary structure of gp41W was determined to be 77% helical, based on basis sets with wavelengths of 178−260 nm, using the SELCON3 CD analysis program (*39*). This is in agreement with the suggestion that gp41W could form an amphipathic helix (*16*), with a preference for the membrane−water interface (*24*) as defined by the Wimley and White membrane interface hydrophobicity scale (*19*).

NMR Spectroscopy. The encouraging results obtained by CD spectroscopy led us to examine the detailed three-dimensional structure of this peptide in a membrane-mimicking environment. Attempts to obtain NMR spectra of gp41W in either SDS-d_{25} or 40% TFE proved to be unsuccessful, resulting in extremely broad resonances in both cases and poor solubility of the peptide in SDS (data not shown). Additionally, the peptide was not very soluble in water alone, giving rise to peptide aggregates and broad chemical shift resonances. The only deuterated detergent commercially available in which the peptide would dissolve, and provide good NMR data, was DPC-d_{38}. After the addition of 100 mM DPC-d_{38}, the peptide appeared to be completely solubilized, and the one-dimensional ^{1}H NMR spectra demonstrated good chemical shift dispersion. DPC was further added to a final concentration of 200 mM to give a detergent:peptide ratio of at least 100:1, ensuring only one peptide molecule per detergent micelle.

The NMR spectra of gp41W in DPC collected at 500 MHz demonstrated that the peptide was adopting a well-defined helical structure, but the resolution at this magnetic field strength was not sufficient to differentiate between many of the chemical shifts. Additionally, the α-protons of Asn and three of the Trp residues had chemical shifts near the water peak (Figure 3), which could not be fully resolved from the water resonance. NOESY and TOCSY spectra were subsequently collected on an 800 MHz NMR spectrometer, which

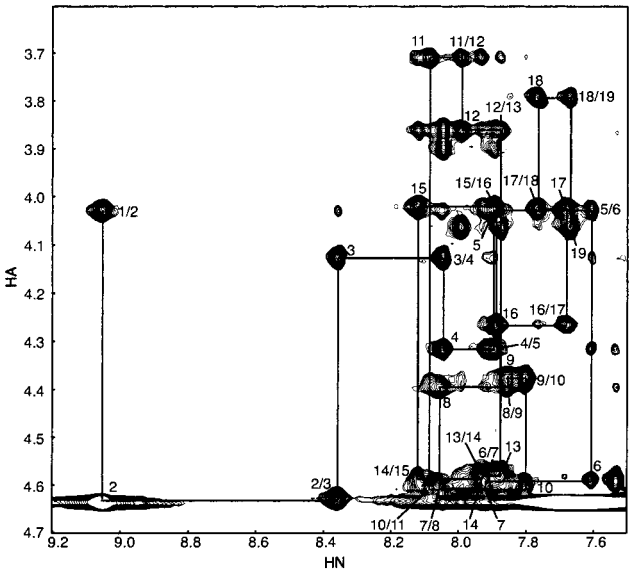

FIGURE 3: Representative section of the 100 ms NOESY spectra of 1.6 mM gp41W in 200 mM DPC-d_{38} micelles (pH 3.51) acquired at 800 MHz. The NH−HA region is shown with the sequential connectivity between NH_i and HA_{i-1} indicated.

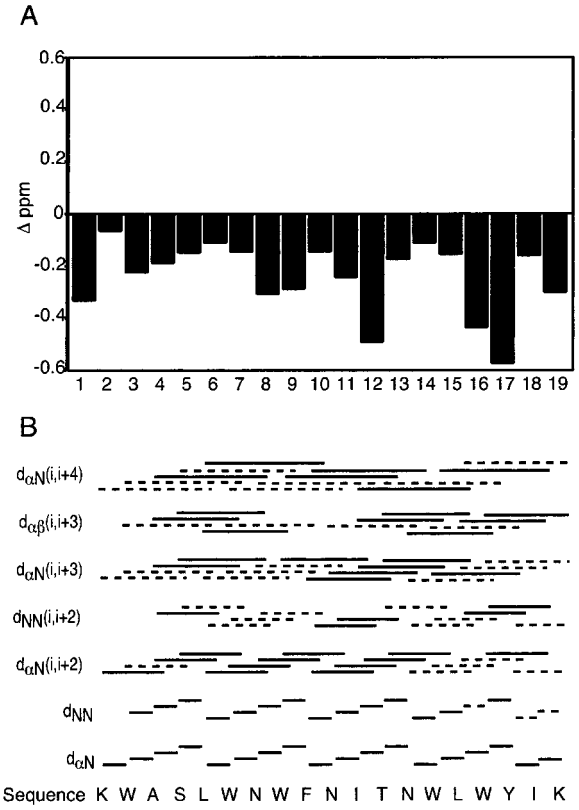

FIGURE 4: (A) Deviation of the HA chemical shifts from random-coil values for each residue of gp41W in DPC micelles, as calculated by NMRView4. (B) NOE connectivities observed in the 100 ms NOESY spectrum, obtained on the 800 MHz NMR spectrometer. Solid lines represent unambiguous connectivities, and dashed lines represent ambiguous connectivities due to spectral overlap.

allowed for the complete assignment of the peptide (Figure 3). Initial examination of the chemical shift distribution of the α-protons suggested that the peptide was adopting a helical structure, which is confirmed by the NOE pattern for the peptide backbone (Figure 4).

As mentioned above, there were many chemical shifts that overlapped in the NMR spectra of gp41W in DPC micelles. While the gp41W peptide has a relatively low molecular mass of 2.6 kDa, binding of the peptide to an approximately 20 kDa DPC micelle increases the correlation time of the peptide (*40*). Therefore, even at the higher resolution provided by the 800 MHz spectrometer, the resonance line widths remained relatively broad, resulting in some chemical shift overlap.

The structures obtained using CNS utilized 435 NOE distance restraints (Table 1). Of these NOEs, those that were possibly ambiguous were given the most logical assignment as determined by the sequence and an assumed helical structure (Figures 2 and 4). To differentiate between the probable contributions for each of the ambiguous NOEs, the automated assignment program ARIA was used. The initial structures obtained with CNS were used as the starting structures for ARIA. Various runs were performed with ARIA to utilize as many unambiguous and ambiguous restraints as possible from both the 800 MHz H_2O and 500 MHz D_2O NOESY spectra. Due to the lower resolution of the D_2O spectrum compared to that of the H_2O spectrum, only those NOEs in the D_2O NOESY that were not observed in the 800 MHz H_2O spectrum were used in the structure calculations. The structures generated using ARIA resulted in NOE restraint files consisting of 525 unambiguous and 69 ambiguous distance restraints (Table 1). While ARIA allows the final structures to undergo a final energy minimization step in a simulated water box, it was decided that due to the presence of the peptide in detergent this step was not applicable.

The results of the deuterium exchange experiment suggested that the amides of the central residues were slowly exchanging (residues 4−18). Signals remained for up to 6 h for most of these amide protons, with all of the amide protons being fully exchanged after 10 h. These results suggest that the amides of the central residues were taking part in hydrogen bonds. While the presence of hydrogen bonds is the most likely explanation, the insertion of a peptide into a detergent micelle will partially segregate the amides from the bulk water, and could slow the amide exchange rate of the more deeply buried amides (*41*). For this reason, no hydrogen bonds were introduced in the structural calculations. It should be noted that although no specific hydrogen bonds were used in the calculations, the amides of the central residues in most of the structures were participating in hydrogen bonds as calculated by MOLMOL (*42*).

NMR Structures. Of the 200 structures generated in the final iteration of ARIA, the 40 lowest-energy structures were retained. The final statistics of these 40 structures are presented in Table 1. gp41W forms a well-defined helix in DPC micelles, with pairwise global rmsds for the 40 structures, fitted to residues 2−18, of 0.47 and 0.81 Å for the backbone and heavy atoms, respectively (Table 1 and Figure 5). As mentioned above, the NOE pattern for gp41W is representative of a helical structure, although the definition of the type of helix that is formed (i.e., α or 3_{10}) is complicated by the presence of both $d_{\alpha N}(i, i+2)$ and $d_{\alpha N}(i, i+4)$ NOEs (Figure 4). NOEs between the $d_{\alpha N}(i, i+2)$ protons are representative of 3_{10} helicity, while $d_{\alpha N}(i, i+4)$ NOEs indicate the presence of α-helical structure (*32*). The variation in the definition of helicity is observed in the 40

Table 1: Structural Statistics for the Final 40 NMR Structures of gp41W in DPC-d_{38} Micelles

	CNS	$ARIA_{input}$	$ARIA_{final}$
no. of distance restraints			
unambiguous NOE	435	544	525
ambiguous NOE	–	165	69
total NOEs	435	709	594
no. of broad dihedral restraints	18	18	18
rms differences from ideal values[a]			
bonds (Å)	$1.37 \times 10^{-3} \pm 7.22 \times 10^{-5}$		
angles (deg)	$3.01 \times 10^{-1} \pm 6.43 \times 10^{-3}$		
impropers (deg)	$1.45 \times 10^{-1} \pm 1.16 \times 10^{-2}$		
van der Waals (kcal/mol)	14.93 ± 0.73		
rms differences from distance restraints[a]			
unambiguous (Å)	$1.25 \times 10^{-2} \pm 6.48 \times 10^{-3}$		
ambiguous (Å)	$(3.16 \pm 2.37) \times 10^{-3}$		
all NOEs (Å)	$1.19 \times 10^{-2} \pm 6.11 \times 10^{-3}$		
nonbonded energies[a]			
electron (kcal/mol)	-68.75 ± 2.44		
van der Waals (kcal/mol)	-21.73 ± 10.76		
pairwise global rmsd (Å)[b]			
backbone	0.47 ± 0.18		
heavy	0.81 ± 0.18		

[a] Calculated by ARIA after eight iterations on the 40 lowest-energy structures. [b] Calculated by MOLMOL for residues 2−18.

structures, with some stretches of the helix forming either α- or 3^{10}-helix, depending on the orientations of the backbone amide protons and carbonyls.

The formation of a helical structure for gp41W causes the Trp residues to form a "collar" around the axial length of the peptide (Figures 5A and 6A), with four of the Trp residues and the single Tyr residue located in the same plane (Figure 6B). The only Trp residue not located in this plane is Trp2, which appears to be below the plane of the other Trp residues (Figure 6B). The location of this residue could be due to an inability to unambiguously assign critical backbone NOEs in the region surrounding Trp2. Specifically, the chemical shift of the α-proton of Trp2 falls directly on the water peak, and even at an NMR frequency of 800 MHz, this peak could not be significantly resolved to obtain much useful structural information. Additionally, the increased disorder of the N-terminal region could also be explained by the weaker hydrogen bonds of the initial three amides, as demonstrated by the D_2O exchange experiments. The only other aromatic residue not in the same plane as the Trp and Tyr residues is Phe9. The greater hydrophobicity of the Phe side chain would likely cause the aromatic ring to be more deeply buried in a detergent micelle compared to the Trp and Tyr residues. This leads us to suggest that it is this face of the helix which is the "bottom" (i.e., the face closest to the hydrocarbon chains of the DPC detergent), with the opposite face more exposed to the solvent (Figure 6).

The polar face of such a structure would be the one opposite that of the Phe-containing face. The polar residues on this face are Lys1, Lys19, Ser4, and Asn7; additionally, the side chain of Asn10 and the oxygen of Thr12 are orientated toward this face (Figure 5C). It should be noted that the side chains of the Lys residues are relatively flexible due to the length of their side chains, resulting in fewer available distance restraints, which could result in variability in their apparent locations. The only polar residue not on this face is Asn13. The burial of the side chain amide of Asn13 in the micelle may be stabilized by hydrogen bonding

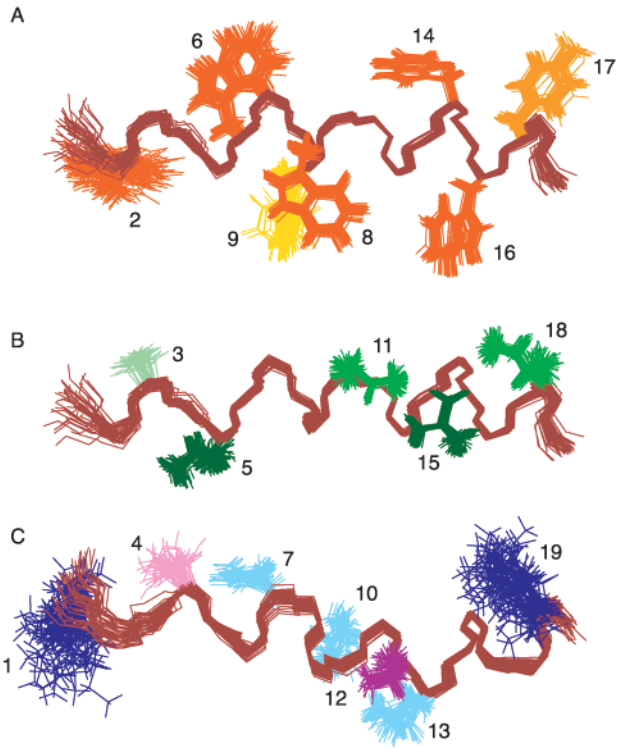

FIGURE 5: Overlay of the 40 lowest-energy structures determined using the iterative assignment program ARIA. Residues 2−18 are overlaid, providing mean pairwise rmsds of 0.47 ± 0.18 and 0.81 ± 0.18 Å for the backbone and heavy atoms, respectively. (A) Aromatic side chains are displayed. Trp residues are dark orange; the Tyr residue is orange, and the Phe residue is gold. (B) Hydrophobic residues are displayed. Leu residues are dark green; Ile residues are green, and the Ala residue is light green. (C) Polar residues are displayed. Lys residues are dark blue, and Asn residues are light blue; the Ser residue is pink, and the Thr residue is purple. The structures in panels A and B are shown in the same orientation, looking down from the "top", while the structures in panel C are rotated approximately 90° to indicate the relative amphipathicity of gp41W in DPC micelles. This figure was created using MOLMOL (*42*).

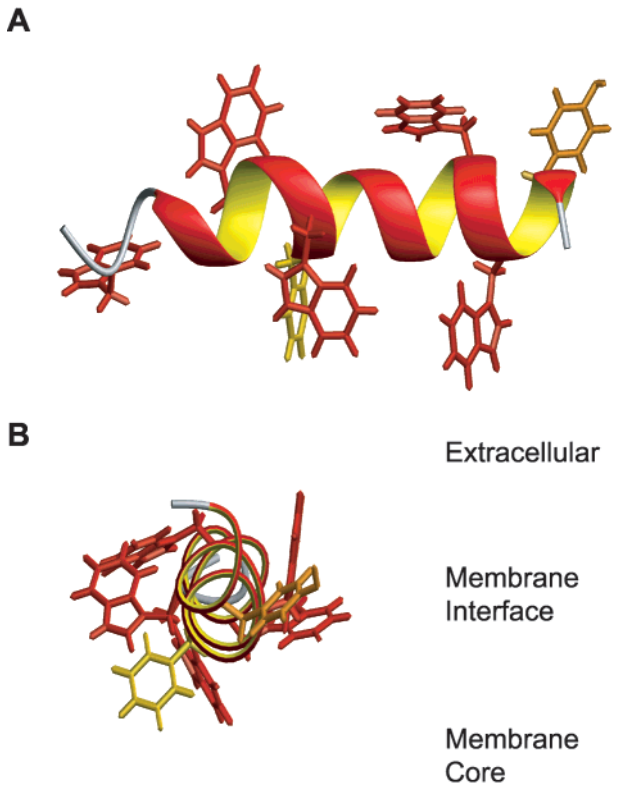

FIGURE 6: Ribbon diagrams of the average gp41W structure. The Trp residues are red; the Tyr residue is gold, and the Phe residue is yellow. (A) View of gp41W, looking down on the polar face of the amphipathic structure from the extracellular space. Four of the Trp residues and the Tyr residue form a collar along the axial length of the peptide. (B) View of gp41W looking down the helical axis, where the polar face is on the top of this structure, and the expected orientation of the peptide in the bilayer is indicated. The four Trp residues and the Tyr residue are within the same plane, which would be oriented within the interfacial region of the bilayer. Trp2 and Phe9 are below the plane of the Trp residues. The burial of Trp2 is probably due to the lack of NOE restraints for sufficiently defining the position of this residue. This figure was created using MOLMOL (*42*).

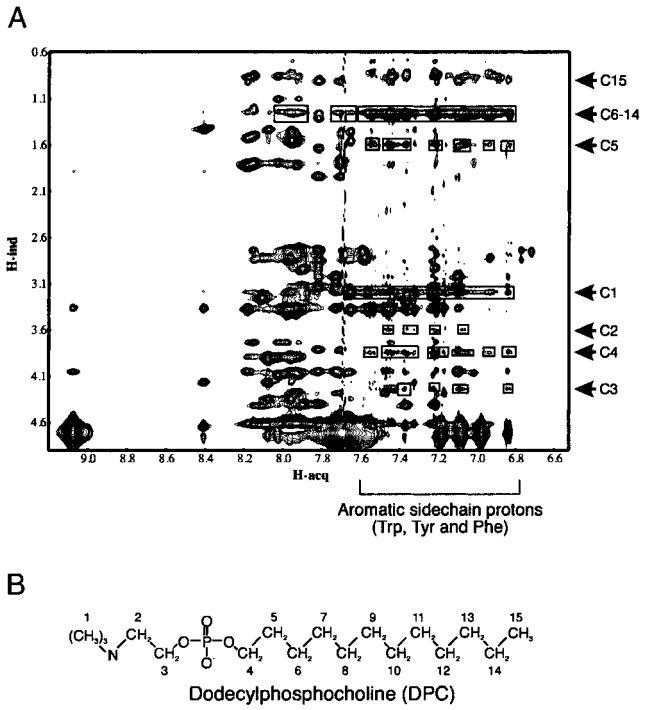

FIGURE 7: (A) NOESY spectra of 2 mM gp41W in 50 mM protonated DPC micelles acquired at 500 MHz with a mixing time of 100 ms. Cross-peaks between protons on DPC and gp41W are indicated by boxes. The chemical shifts corresponding to the protons of each of the carbons in DPC are indicated on the *y*-axis. The majority of the additional cross-peaks observed in the presence of [^{1}H]DPC result from NOEs to the aromatic protons of Trp, Tyr, and Phe, but due to the lack of significant resolution, these peaks cannot be unambiguously assigned. (B) Numbering scheme of the protons of DPC.

with the backbone oxygen of Phe10. The presence of the bulk of the polar residues in a single face supports the above suggestion that this is the more solvent exposed face when the peptide is inserted in a DPC micelle. The remaining hydrophobic residues form the core of the peptide (Figure 5B). While this peptide does appear to be relatively amphipathic, some of the hydrophobic side chains are on the same face as the more polar residues. This does not appear to be easily avoidable, due to the shortage of other polar residues in the peptide that could be positioned on this face.

To confirm the interfacial localization of the gp41W peptide in DPC micelles, NOESY spectra of the peptide in [^{1}H]DPC micelles were acquired for observation of intermolecular NOEs between the peptide and the DPC detergent (Figure 7). This technique has been used previously to determine the location and orientation of nisin in both SDS and DPC micelles (*43*). The ratio of DPC to peptide was decreased to 25:1 to intensify any NOEs between the peptide and detergent as was done for nisin in DPC micelles (*43*). The lack of any significant changes in the gp41W chemical shifts in the presence of [^{1}H]DPC, compared to the shifts in the presence of perdeuterated DPC, suggests that the structure was not affected by the decreased ratio of detergent to peptide. The assignment of the DPC proton resonances was adapted from the published chemical shifts (*43*) and confirmed via standard NMR techniques. Nearly all of the aromatic protons of gp41W appear to have NOEs to the interfacial protons of DPC, while there are very few NOEs between DPC and the amides of the gp41W peptide (Figure 7). This is to be expected since in the NMR structures the aromatic rings form a ring around the peptide blocking the amide protons from interacting with the protons of DPC. While there are also NOEs to the methylene carbons of the acyl chain of DPC, it is not possible to resolve the protons of the nine central carbons, which inhibits the accurate determination of the depth of the peptide in the micelle. It should be noted that there are no NOE cross-peaks between the peptide and the buried terminal methyl group of DPC (protons of C15) (Figure 7), as no new peaks were visible in this region compared to that of the peptide in perdeuterated DPC. This confirms that the peptide does not pass through the micelle, but instead is localized to the surface. Unfortunately, due to the very intense signals of the protonated DPC, it was not possible to reliably observe any NOEs between [^{1}H]DPC and the aliphatic protons of the peptide in the upfield region of the spectrum.

DISCUSSION

The fusion peptide of gp41 was determined to be the glycine-rich peptide at the N-terminus of gp41 by sequence similarities with the influenza fusion peptide (*44*). Using NMR spectroscopy, Chang et al. (*45*) observed that the fusion

peptide forms an α-helix in SDS micelles. The ability of the fusion peptide to interact with and cause disruption of both phospholipid vesicles and cellular membranes has been extensively studied (*24*, *46*−*51*). Recently, the membrane-proximal Trp-rich region of gp41 has been suggested to aid in the disruption of membranes during the gp41-mediated fusion process (*16*), with a potentially higher affinity for membrane disruption than the N-terminal fusion peptide (*24*). The disruption of phospholipid membranes by tryptophan-containing peptides is reminiscent of the Trp-rich class of antimicrobial peptides, specifically indolicidin and tritrpticin, which have been shown to bind to and disrupt both bacterial membranes and vesicles (refs *52*−*55* and unpublished observations).

Our inability to obtain NMR spectra of gp41W under conditions other than DPC micelles suggests that the peptide is selective for a zwitterionic membrane-like environment, similar to what is found in the mammalian extracellular membrane. The broad NMR resonances and aggregation observed in H_2O and TFE are indicative of gp41W forming multimers of varying size, as has been observed for the related HIV DP178 peptide (*56*). While SDS is routinely used as a membrane mimetic (*40*), the negatively charged headgroup of this detergent is unlike a mammalian cellular membrane. The major difference between SDS and DPC is the replacement of the negative sulfate moiety of SDS with a zwitterionic choline headgroup in the DPC molecule. It is likely that the presence of this zwitterionic group is responsible for the solubility and excellent NMR spectra that were obtained for the gp41W peptide. It has been reported that, while a DPC micelle does not possess the glycerol portion of a phospholipid, DPC does possess some of the characteristics of phosphatidylcholine lipids (*57*).

The helical structure of gp41W in membrane-mimetic DPC micelles described here confirms the previous suggestion that this sequence forms a helix when interacting with phospholipid bilayers (*16*) and suggests a likely orientation of the peptide in membranes. The majority of the NOEs that were observed between protonated DPC and the gp41W peptide were to the aromatic rings of the peptide and not to the amide protons. The lack of NOEs to the terminal methyl of the DPC detergent confirms that this peptide does not become inserted through the core of the micelle, but instead is "floating" at the H_2O−DPC interface of the micelle. The indole ring moiety of tryptophan has a preference for being positioned at the membrane−water interfacial region (*18*−*22*). The collar of tryptophan side chains (Figure 6), with four of the five indole rings, as well as the tyrosine side chain within the same plane, indicates that gp41W inserts perpendicular to the bilayer normal, near the membrane interfacial region (Figure 8). As mentioned above, such an orientation positions the bulk of the polar residues on the same face of the helix, creating an amphipathic structure.

The collar of aromatic residues appears to have been optimized to position the maximum number of Trp residues along the length of the helix (Figure 6A). This suggests that there was an evolutionary constraint for the incorporation of such a relatively rare aromatic residue. The presence of multiple aromatic groups immediately adjacent to the transmembrane helix of the fusion protein is found not only in HIV but also in many other viral envelope proteins, including

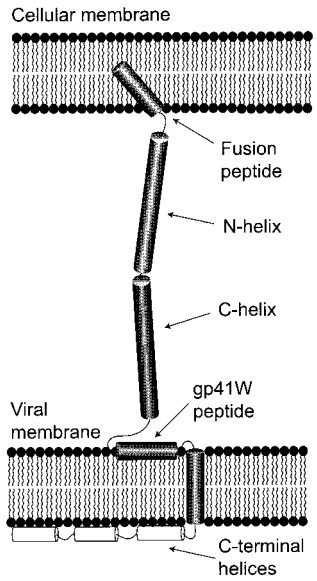

FIGURE 8: Possible model for the location of the gp41W peptide in the membrane following the conformational change to the prehairpin intermediate induced by CD4 binding. The gp41W peptide would become attached to the surface of the viral membrane as an α-helix, while the fusion peptide becomes inserted into the target cellular membrane. Only a monomer is shown of the gp41 trimer complex for simplicity.

SIV gp41, the Ebola G protein, the vesicular stomatitis virus (VSV) G protein, and the MoMuLV transmembrane protein (*16*, *24*, *58*). The Trp-rich region of the VSV G protein has been demonstrated to be essential for efficient budding of the virus from the membrane (*58*). The association of such aromatic-rich regions with various membrane interacting processes demonstrates a theme, supporting a role for the Trp-rich segments in virus assembly and membrane fusion during infection.

The preference of the indole ring of tryptophan for being inserted into the membrane−water interface is a common theme in membrane proteins (*59*). Additionally, we have recently observed that several Trp-rich cationic antimicrobial peptides of varying lengths will be positioned at the membrane−water interface (unpublished observations). The relatively bulky side chain of tryptophan and its location near the glycerol portion of the phospholipid leaflet may sterically disturb the organization of phospholipids, leading to membrane disruption. Robison et al. (*58*) have suggested that the insertion of the Trp-rich membrane-proximal region of the VSV G protein may induce positive curvature in the viral membrane during viral budding, instigating membrane disruption. This is similar to the action of many antimicrobial peptides, which will insert into the interfacial region of phospholipid bilayers, leading to the induction of positive curvature strain (*60*, *61*). The positioning of gp41W in the viral membrane following the conformational change of gp41 to the "prehairpin" intermediate (Figure 8) may disrupt the viral membrane in a manner similar to the proposed mechanism of antimicrobial peptides.

The gp120−gp41 conformational change observed upon receptor binding causes an increase in the hydrophobic surface area that is accessible to the solvent (*8*). As mentioned above, the gp41W peptide forms a partially amphipathic structure, yet a few of the peptide's hydrophobic residues are also located on the "polar" face of the structure.

The exposure of this partially hydrophobic and/or polar face during the gp41 conformational change could, in part, contribute to this increase in hydrophobicity. Bentz (*62*) recently proposed that the insertion of one or more fusion peptides from influenza HA into the viral membrane could create a hydrophobic defect in the membrane, leading to the disruption of the viral membrane. The Trp-rich peptide of gp41 may play a similar role in aiding the formation of a hydrophobic defect. The use of gp41W instead of the gp41 fusion peptide to disrupt the viral membrane could actually decrease the likelihood of the HIV virus inactivating itself by an irreversible insertion of the fusion peptide into the viral membrane.

The N-terminal end of gp41W (residues 665−673) overlaps with the HIV inhibitory peptide DP178 (residues 638−673) (*25*). The N-terminus of the DP178 peptide can bind selectively to the fusion peptide of gp41 (*63*−*65*), suggesting a possible mechanism for the inhibition of the action of the fusion peptide and gp41W in the native metastable nonfusogenic conformation prior to the initiation of the fusion reaction. Following the gp120−CD4 interaction, the conformational change to the prehairpin intermediate would release both the fusion peptide and gp41W, allowing them to interact with their respective membrane targets. DP178 has also been shown to block the gp41 conformational change, thought to form the "fusion-active" conformation (*66*), supposedly by binding to the N-helix (N51, residues 540−589) of the prehairpin intermediate prior to the formation of the coiled-coil hairpin observed in the X-ray structure (*67*). In addition, DP178 can clamp HIV fusion in a lipid mixing state (*68*), suggesting the viral and cellular membranes are in a state of hemifusion prior to the final conformational change. A recent study has suggested that DP178 may inhibit membrane fusion via a second mode of inhibition, through an interaction with the complementary region of gp41 positioned within the membrane (Figure 8), inhibiting oligomerization of gp41 (*69*). The attachment of the DP178 peptide to the membrane surface appears to be mediated by the C-terminal Trp-rich region of DP178, the region that overlaps with the gp41W peptide.

The confirmation here that the gp41W sequence adopts a helical structure in the presence of membrane-mimetic DPC micelles supports its role in membrane fusion through a "Velcro-like" interaction with the phospholipid bilayer. Detailed examination of the interactions between the gp41W sequence and the remainder of the gp41 protein and their interactions with membranes is required for a fuller understanding of the mechanism of HIV gp41-mediated membrane fusion.

ACKNOWLEDGMENT

We thank the Canadian National High Field NMR Centre (NANUC) for their assistance and use of the facilities. Operation of NANUC is funded by the Canadian Institutes of Health Research, the Natural Science and Engineering Research Council of Canada, and The University of Alberta. We are indebted to Mr. K. Islam (University of Pittsburgh School of Medicine) for peptide synthesis and Mr. A. Weljie (University of Calgary) for discussions on the use of the ARIA program. D.J.S. thanks Dr. P. Mobley for insightful discussions on the HIV mechanism. We thank Dr. C. Hastings and Dr. K. Gardner for making available various NMRView scripts.

SUPPORTING INFORMATION AVAILABLE

A table of the chemical shift assignment of the gp41W peptide in DPC micelles and a Ramachandran plot. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hernandez, L. D., Hoffman, L. R., Wolfsberg, T. G., and White, J. M. (1996) *Annu. Rev. Cell Dev. Biol. 12*, 627−661.
2. Freed, E. O., and Martin, M. A. (1995) *J. Biol. Chem. 270*, 23883−23886.
3. Dalgleish, A. G., Beverly, P. C. L., and Clapham, P. R. (1984) *Nature 312*, 763−767.
4. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature 393*, 648−659.
5. Feng, Y., Broder, C. C., Kennedy, P. E., and Berger, E. A. (1996) *Science 272*, 872−877.
6. Sattentau, Q. J., Moore, J. P., Vignaux, F., Traincard, F., and Poignard, P. (1993) *J. Virol. 67*, 7383−7393.
7. Matthews, T. J., Wild, C., Chen, C. H., Bolognesi, D. P., and Greenberg, M. L. (1994) *Immunol. Rev. 140*, 93−104.
8. Jones, P. L., Korte, T., and Blumenthal, R. (1998) *J. Biol. Chem. 273*, 404−409.
9. Chan, D. C., and Kim, P. S. (1998) *Cell 93*, 681−684.
10. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature 387*, 426−430.
11. Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981) *Nature 289*, 366−373.
12. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell 89*, 263−273.
13. Tan, K., Liu, J.-H., Wang, J.-H., Shen, S., and Lu, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12303−12308.
14. Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) *Nature 371*, 37−43.
15. Caffrey, M., Cai, M., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) *EMBO J. 17*, 4572−4584.
16. Salzwedel, K., West, J. T., and Hunter, E. (1999) *J. Virol. 73*, 2469−2480.
17. Munoz-Barroso, I., Salzwedel, K., Hunter, E., and Blumenthal, R. (1999) *J. Virol. 73*, 6089−6092.
18. Kachel, K., Asuncion-Punzalan, E., and London, E. (1995) *Biochemistry 34*, 15475−15479.
19. Wimley, W. C., and White, S. H. (1996) *Nat. Struct. Biol. 3*, 842−848.
20. Persson, S., Killian, J. A., and Lindblom, G. (1998) *Biophys. J. 75*, 1365−1371.
21. Yau, W., Wimley, W. C., Gawrisch, K., and White, S. H. (1998) *Biochemistry 37*, 14713−14718.
22. de Planque, M. R. R., Kruijtzer, J. A. W., Liskamp, R. M. J., Marsh, D., Greathouse, D. V., Koeppe, R. E., de Kruijff, B., and Killian, J. A. (1999) *J. Biol. Chem. 274*, 20839−20846.
23. Suarez, T., Nir, S., Goni, F. M., Saez-Cirion, A., and Nieva, J. L. (2000) *FEBS Lett. 477*, 145−149.
24. Suarez, T., Gallaher, W. R., Agirre, A., Goni, F. M., and Nieva, J. L. (2000) *J. Virol. 74*, 8038−8047.
25. Wild, C., Greenwell, T., and Matthews, T. (1993) *AIDS Res. Hum. Retroviruses 9*, 1051−1053.
26. Muster, T., Steindl, F., Purtscher, M., Trkola, A., Klima, A., Himmler, G., Rüker, F., and Katinger, H. (1993) *J. Virol. 67*, 6642−6647.
27. Tencza, S. B., Douglass, J. P., Creighton, D. J., Jr., Montelaro, R. C., and Mietzner, T. A. (1997) *Antimicrob. Agents Chemother. 41*, 2394−2398.
28. Callihan, D., West, J., Kumar, S., Schweitzer, B. I., and Logan, T. L. (1996) *J. Magn. Reson. 112*, 82−85.
29. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. 102A*, 241−245.

30. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
31. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
32. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
33. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
34. Nilges, M. (1995) *J. Mol. Biol. 245*, 645−660.
35. Nilges, M. (1997) *Folding Des. 2*, 53−57.
36. Nilges, M., and O'Donoghue, S. (1998) *Prog. NMR Spectrosc. 32*, 107−139.
37. Linge, J., and Nilges, M. (1999) *J. Biomol. NMR 13*, 51−59.
38. Woody, R. W. (1994) *Eur. Biophys. J. 23*, 253−262.
39. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999) *Protein Sci. 8*, 370−380.
40. Henry, G. D., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 515−535.
41. Spyracopoulos, L., and O'Neil, J. D. J. (1994) *J. Am. Chem. Soc. 116*, 1395−1402.
42. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
43. van den Hooven, H. W., Spronk, C. A. E. M., van de Kamp, M., Konings, R. N. H., Hilbers, C. W., and van de Ven, F. J. M. (1996) *Eur. J. Biochem. 235*, 394−403.
44. Gallaher, W. R. (1987) *Cell 50*, 327−328.
45. Chang, D.-K., Cheng, S.-F., and Chien, W.-J. (1997) *J. Virol. 71*, 6593−6602.
46. Nieva, J. L., Nir, S., Muga, A., Gopni, F. M., and Wilschut, J. (1994) *Biochemistry 33*, 3201−3209.
47. Mobley, P. W., Lee, H. F., Curtain, C. C., Kirkpatrick, A., Waring, A. J., and Gordon, L. M. (1995) *Biochim. Biophys. Acta 1271*, 304−314.
48. Kliger, Y., Aharoni, A., Rapaport, D., Jones, P., and Blumenthal, R. (1997) *J. Biol. Chem. 272*, 13496−13505.
49. Mobley, P. W., Waring, A. J., Sherman, M. A., and Gordon, L. M. (1999) *Biochim. Biophys. Acta 1418*, 1−18.
50. Agirre, A., Fach, C., Goni, F. M., Medelsohn, R., Valpuesta, J. M., Wu, F., and Nieva, J. L. (2000) *Biochim. Biophys. Acta 1467*, 153−164.
51. Durell, S. R., Martin, I., Ruysschaert, J.-M., Shai, Y., and Blumenthal, R. (1997) *Mol. Membr. Biol. 14*, 97−112.
52. Selsted, M. E., Novotny, M. J., Morris, W. L., Tang, Y.-Q., Smith, W., and Cullor, J. S. (1992) *J. Biol. Chem. 267*, 4292−4295.
53. Ladohkhin, A. S., Selsted, M. E., and White, S. H. (1997) *Biophys. J. 72*, 794−805.
54. Lawyer, C., Pai, S., Watabe, M., Borgia, P., Mashimo, T., Eagleton, L., and Watabe, K. (1996) *FEBS Lett. 390*, 95−98.
55. Schibli, D. J., Hwang, P. M., and Vogel, H. J. (1999) *Biochemistry 38*, 16749−16755.
56. Lawless, M. K., Barney, S., Guthrie, K. I., Bucy, T. B., Petteway, S. R., Jr., and Merutka, G. (1996) *Biochemistry 35*, 13697−13708.
57. Beswick, V., Guerois, R., Cordier-Ochsenbein, F., Coïc, Y.-M., Huynh-Dinh, T., Tostain, J., Noël, J.-P., Sanson, A., and Neumann, J.-M. (1998) *Eur. Biophys. J. 28*, 48−58.
58. Robison, C. T., and Whitt, M. A. (2000) *J. Virol. 74*, 2239−2246.
59. Reithmeier, R. A. F. (1995) *Curr. Opin. Struct. Biol. 5*, 491−500.
60. Matsuzaki, K., Sugishita, K., Ishibe, N., Ueha, M., Nakata, S., Miyajima, K., and Epand, R. M. (1998) *Biochemistry 37*, 11856−11863.
61. Epand, R. M., and Vogel, H. J. (1999) *Biochim. Biophys. Acta 1462*, 11−28.
62. Bentz, J. (2000) *Biophys. J. 78*, 886−900.
63. Jiang, S., Lin, K., Strick, N., and Neurath, A. R. (1993) *Biochem. Biophys. Res. Commun. 195*, 533−538.
64. Neurath, A. R., Lin, K., Strick, N., and Jiang, S. (1995) *AIDS Res. Hum. Retroviruses 11*, 189−190.
65. Mobley, P. W., Pilpa, R., Brown, C., Waring, A. J., and Gordon, L. M. (2001) *AIDS Res. Hum. Retroviruses 17*, 311−327.
66. Furuta, R. A., Wild, C. T., Wng, Y., and Weiss, C. D. (1998) *Nat. Struct. Biol. 5*, 276−279.
67. Kilger, Y., and Shai, Y. (2000) *J. Mol. Biol. 295*, 163−168.
68. Muñoz-Barroso, I., Durell, S., Sakaguchi, K., Apella, E., and Blumenthal, R. (1998) *J. Cell Biol. 140*, 315−323.
69. Kliger, Y., Gallo, S. A., Peisajovich, S. G., Muñoz-Barroso, I., Avkin, S., Blumenthal, R., and Shai, Y. (2000) *J. Biol. Chem. 276*, 1391−1397.

BI010640U