A Monomeric 3_{10}-Helix Is Formed in Water by a 13-Residue Peptide Representing the Neutralizing Determinant of HIV-1 on gp41[†,‡]

Zohar Biron,[§] Sanjay Khare,[||] Abraham O. Samson,[§] Yehezkiel Hayek,[§] Fred Naider,[||] and Jacob Anglister*[,§]

Department of Structural Biology, The Weizmann Institute of Science, Rehovot 76100, Israel, and Department of Chemistry, College of Staten Island and Graduate Center of the City University of New York, 2800 Victory Boulevard, Staten Island, New York 10314

Received June 7, 2002; Revised Manuscript Received August 4, 2002

ABSTRACT: The peptide $gp41_{659-671}$ (ELLELDKWASLWN) comprises the entire epitope for one of the three known antibodies capable of neutralizing a broad spectrum of primary HIV-1 isolates and is the only such epitope that is sequential. Here we present the NMR structure of $gp41_{659-671}$ in water. This peptide forms a monomeric 3_{10}-helix stabilized by $i,i+3$ side chain−side chain interactions favored by its primary sequence. In this conformation the peptide presents an exposed surface, which is mostly hydrophobic and consists of conserved HIV-1 residues. The presence of the 3_{10}-helix is confirmed by its characteristic CD pattern. Studies of the 3_{10}-helix have been hampered by the absence of a model peptide adopting this conformation. $gp41_{659-671}$ can serve as such a model to investigate the spectral characteristics of the 3_{10}-helix, the factors that influence its stability, and the propensity of different amino acids to form a 3_{10}-helix. The observation that the 3_{10}-helical conformation is highly populated in the peptide $gp41_{659-671}$ indicates that the corresponding segment in the cognate protein is an autonomous folding unit. As such, it is very likely that the helical conformation is maintained in gp41 throughout the different tertiary structures of the envelope protein that form during the process of viral fusion. However, the exposure of the $gp41_{659-671}$ segment may vary, leading to changes in the reactivity of anti-gp41 antibodies in the different stages of viral fusion. Since $gp41_{659-671}$ is an autonomous folding unit, peptide immunogens consisting of the complete $gp41_{659-671}$ sequence are likely to induce antibodies highly cross-reactive with HIV-1.

The protein gp160 of the human immunodeficiency virus type 1 (HIV-1)[1] is the precursor of the surface glycoproteins gp120 and gp41. The latter is a transmembrane protein that mediates the fusion of the virus with host cells. The three-dimensional structure of the HIV-1 gp41 core, made of two segments containing heptad leucine/isoleucine repeats, N36 ($gp41_{546-581}$) and C34 ($gp41_{628-661}$), was solved by X-ray crystallography (*1*−*3*), and the structure of the SIV gp41 core was solved by multidimensional NMR (*4*). The gp41 core is a six-helix bundle that exhibits structural similarity to other fusion proteins such as the hemagglutinin fusion protein of the influenza virus (*5*) and is assumed to be the fusogenic or postfusogenic form of this protein (*6*). Antibodies against this structure do not seem to neutralize the virus possibly because once the six-helical coiled-coil structure is formed in the fusion process, gp41 is not accessible to HIV-1 neutralizing antibodies (*7*).

The C-terminal region (residues 660−683) of the gp41 ectodomain preceding the membrane-spanning domain is rich in tryptophan residues that are conserved in lentiviruses (*8*). Deletion of residues 660−670 (LLELDKWASLW) results in partial dissociation of the oligomeric structure of gp41 (*9*), abolishes fusion, and decreases gp160 precursor cleavage. Mutation of the first three tryptophan residues (W666, W670, and W672) suffices to eliminate viral fusion (*8*).

The C-terminal tryptophan-rich region overlaps the peptide T-20 ($gp41_{638-673}$), which is a strong inhibitor of HIV-1 fusion with an EC_{50} of 0.5 ng/mL (*10*). This peptide is presently in a phase III clinical trial (*11*). In contrast to all currently available anti-HIV drugs that inhibit the virus in already infected cells, T-20 is expected to represent a new class of anti HIV-1 drugs that inhibit the fusion of the virus with its target cells, thus preventing infection of new cells. A peptide corresponding to the C-terminal residues 661−673 of T-20 inhibits HIV-1 infection with an EC_{50} of 20 μg mL^{-1}.

[†] This study was supported by National Institutes of Health Grants R01 GM53329 (to J.A.) and GM22086 (to F.N.). J.A. is the Joseph and Ruth Owades Professor in Chemistry. F.N. is the Leonard and Esther Kurtz Term Professor at the College of Staten Island.

[‡] The NMR data and the atomic coordinates of the 25 lowest energy structures (1LCX) and the average structure (1LB0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

* Corresponding author. Phone: 011-972-8-934-3394. Fax: 011-972-8-934-4136. E-mail: Jacob.Anglister@weizmann.ac.il.

[§] The Weizmann Institute of Science.

[||] College of Staten Island and the Graduate Center of CUNY.

[1] Abbreviations: Ac, acetyl; Ac-$gp41_{659-671}$-NH_2, $gp41_{659-671}$ acetylated at the N-terminus and amidated at the C-terminus; CD, circular dichroism; DQF-COSY, double-quantum-filtered correlation spectroscopy; EC_{50}, concentration that causes 50% inhibition of the virus; gp, glycoprotein; $gp41_{659-671}$, a gp41 peptide comprising residues 659−671; HIV-1, human immunodeficiency virus type 1; NOESY, nuclear Overhauser effect spectroscopy; SIV, simian immunodeficiency virus.

10.1021/bi026261y CCC: $22.00
Published on Web 09/27/2002

The tryptophan-rich domain of gp41 includes also the ELDKWA sequence that is the core epitope for the HIV-1 neutralizing monoclonal antibody 2F5 (*12*). MALDI-MS protection assay suggested that the complete 2F5 epitope consists of 16 residues mapped to $gp41_{656-671}$ (*13*). The 2F5 antibody has a 5-fold higher affinity to the $gp41_{659-671}$ peptide in comparison to its affinity to gp41 [K_a of 1.3×10^9 M^{-1} in comparison with 2.4×10^8 M^{-1} (*14*)], indicating that $gp41_{659-671}$ contains all residues important for 2F5 binding. Only three antibodies, 2F5, 2G12, and G1b12, were found to be capable of neutralizing a broad spectrum of primary HIV-1 isolates (*15*). Passive immunization with these antibodies and especially their combination protect macaques against a challenge of a pathogenic HIV-1/SIV chimeric virus (*16*, *17*). Immunogens that present the epitopes of these three antibodies in a way that mimics their structure on native HIV-1 may be an important component of an anti-HIV-1 vaccine (*7*, *13*, *18*).

The antibody 2F5 binds the gp120/gp41 complex, and its affinity diminishes after CD4 binding to gp120 (*19*). It has been shown recently that 2F5 binding to the peptide C43 ($gp41_{624-666}$) is inhibited by preincubation of C43 with the peptide N51 ($gp41_{540-590}$) as a result of complex formation between these two peptides (*20*). These findings indicate that the 2F5 epitope is exposed in the prefusogenic form of gp41 (before CD4 binding to gp120) and is occluded or disrupted in the fusogenic form of gp41 (after CD4 binding to gp120). The structure of gp41 in the prefusogenic form, which is the structure relevant for vaccine development, is not known.

The 3_{10}-helix is the fourth most common secondary structural motif in proteins. Three to four percent of residues in crystal structures of proteins are in 3_{10}-helices, accounting for 20% of the observed helices (*21*) in these biopolymers. Due to a decreased stability of the 3_{10}-helices in comparison to α-helices, segments in a 3_{10}-helical conformation are usually very short and on average only three to four residues long (*21*). The 3_{10}-helix has recently attracted the attention of biochemists and structural biologists, and it has been suggested that it may act as a folding intermediate in α-helix formation (*22*, *23*). The first water-soluble 3_{10}-helical peptide was reported recently, but it was composed only of unnatural amino acids (*24*), and five C^{α}-tetrasubstituted α-amino acids stabilized the 3_{10}-helix conformation in this heptapeptide. These unnatural amino acids lack α protons necessary for NMR detection of the $\alpha N(i,i+2)$ and $\alpha N(i,i+4)$ NOEs that differentiate a 3_{10}-helix from an α-helix and the $\alpha N(i,i+3)$ and $\alpha\beta(i,i+3)$ NOEs that characterize both helices and which can be used to estimate the relative populations of 3_{10}- and α-helices (*23*, *25*, *26*).

Here we report that a nine residue long 3_{10}-helix is formed in water by a monomeric tridecapeptide corresponding to residues 659−671 (ELLELDKWASLWN) of the envelope glycoprotein gp41 of HIV-1 ($gp41_{659-671}$). The $gp41_{659-671}$ peptide contains the entire epitope of the HIV-1 neutralizing antibody 2F5. This peptide can be used to characterize the spectral properties of 3_{10}-helices and to investigate the parameters that stabilize this type of helical conformation. Since $gp41_{659-671}$ is an autonomous folding unit, it most likely represents the structure of the corresponding segment in gp41, and therefore immunogens consisting of the complete $gp41_{659-671}$ sequence are likely to induce antibodies highly cross-reactive with HIV-1.

EXPERIMENTAL PROCEDURES

Peptide Synthesis. The peptides ELLELDKWASLWN ($gp41_{659-671}$) and Ac-ELLELDKWASLWN-NH_2 (Ac-$gp41_{659-671}$-NH_2) were synthesized on an Applied Biosystems peptide synthesizer. The crude peptides were purified on a Vydac polymer column using an acetonitrile gradient. The water and acetonitrile contained 0.1% NH_3. The peptide sequences were verified by mass spectrometry and amino acid analysis.

Sample Preparation. NMR samples were prepared by dissolving the peptide in a 50 mM aqueous solution of CD_3-$COONH_4$, pH 7.7, containing 0.05% NaN_3. The 95% H_2O/5% D_2O contained 1.85 mM peptide, and the 99.9% D_2O solution contained 3 mM peptide.

Circular Dichroism Studies. The CD spectra of the peptide were recorded on an AVIV model 62 DS CD instrument (AVIV Associates, Lakewood, NJ) at 277, 283, and 298 K. Data were scanned at 1 nm intervals, and five spectra were averaged for each condition. Peptide concentrations in solutions were determined from UV measurements at 280 nm using an extinction coefficient of 7000 M^{-1} cm^{-1} per Trp residue. CD intensities are expressed as mean residue ellipticities (deg·cm^2 $dmol^{-1}$).

Diffusion Coefficient Measurements. A longitudinal encode−decode (LED) experiment (*27*) was used for measurements of translational self-diffusion coefficients on a Bruker DMX400 spectrometer using a 5 mm triple resonance probe with a single gradient (Z). The diffusion coefficient was measured in triplicate at two concentrations of the $gp41_{659-671}$ and Ac-$gp41_{659-671}$-NH_2 peptides (3 mM, 0.1 mM) in 50 mM ammonium acetate buffer, pH 7.7, in D_2O. Diffusion coefficients were measured with 3 ms duration of the field gradient and 100 ms delay between the gradients. Measurements were carried out at 277 K using 62 G/cm Z-gradient. A series of 10 spectra were recorded in 2D mode for each measurement.

Sedimentation Equilibrium Measurement. Sedimentation equilibrium measurements were conducted in a Beckman XL-A analytical ultracentrifuge. Sample concentration was limited to 0.1 mM due to the high extinction coefficient of the peptide. The buffer and the pH conditions were identical to those in the NMR and the diffusion measurements. Equilibrium data were collected at 42000 rpm at 277 K. Concentration gradients were monitored by the absorbance at 280 nm. Equilibrium had been reached after 5 days.

NMR Measurements. NMR spectra were recorded on Bruker DMX 500 MHz and DRX 800 MHz spectrometers (Bruker). Measurements were conducted at 277 K. NOESY (*28*) measurements in H_2O used a water-flip back pulse and a 3-9-19 WATERGATE (water suppression by gradient tailored excitation) pulse sequence with gradients for water suppression (*29*). $^3J_{HN\alpha}$ coupling constants were measured from a 1D spectrum. For each residue the amide proton T_2 relaxation time was determined using a 1,1-echo sequence.

Stereospecific Assignment. For stereospecific assignment two additional NOESY spectra were recorded in H_2O and D_2O solutions with a short mixing time of 80 ms, and the intensities of the H^N/H^{β} and H^{α}/H^{β} cross-peaks were compared. A DQF-COSY spectrum in D_2O was recorded under the same conditions to obtain $^3J_{H\alpha H\beta}$ couplings. The data from these three experiments were used to determine

Table 1: NMR Constraints and Structural Statistics for 25 Structures of $gp41_{659-671}$ Peptide

distance constraints	202
hydrogen bond constraints	1
ϕ dihedral angle constraints	9
χ_1 dihedral angle constraints	4
maximum individual NOE violation (Å)	0.50
rmsd of NOE violations	0.02 ± 0.001
deviations from ideal covalent geometry	
bonds lengths (Å)	0.0014 ± 0.0001
bond angles (deg)	0.4727 ± 0.0001
improper angles (deg)	0.0970 ± 0.0001
mean rmsd values (Å)	
backbone atoms, residues 661−669	0.82
all heavy atoms, residues 661−669	1.98

the χ_1 dihedral angles (*30*) of residues E662, W666, S668, and W670.

Experimental Constraints. NOESY spectra were acquired using a 250 ms mixing time. The HN(i)/Hα(i+1) NOE was used for distance calibration. NOE-derived distances (202) were constrained between 1.8 Å and 130% of the calculated distance (*30*). The ϕ dihedral angles along the polypeptide backbone were constrained between $-62 \pm 25°$ (*30*) for all of the amino acids exhibiting $^3J_{HN\alpha}$ <5.6 Hz, resulting in nine dihedral angle constraints. For E662, W666, and W670 the χ_1 dihedral angle was constrained to $180 \pm 30°$ and for S668 to $60 \pm 30°$.

Structure Calculations. The structure of $gp41_{659-671}$ was calculated by the CNS program (Crystallography and NMR System) using the distance geometry−simulated annealing method (*31*), starting from the extended strand conformation (*32*). Initially no hydrogen bonds were input as constraints. However, a short i,i+3 distance between the carbonyl oxygen of D664 and the amide proton of A667, exhibited in all accepted structures, was included as a hydrogen bond constraint in the final structure calculations. Ninety-six structures with no violations of the NMR constraints were generated (out of 100 calculated structures), and the 25 structures with the lowest energy were chosen for statistical analyses (Table 1).

RESULTS

Estimation of Molecular Mass of $gp41_{659-671}$ in Water. The molecular mass of $gp41_{659-671}$ was determined using sedimentation equilibrium at 277 K to find out whether the peptide forms oligomers. Due to the high extinction coefficient at 280 nm, the molecular mass was determined using a 0.1 mM solution and was found to be 1789 Da, in good agreement with the expected mass of 1613 Da. The translational diffusion coefficient of the peptide in D_2O, determined by NMR, was found to be $(1.37 \pm 0.03) \times 10^{-6}$ and $(1.35 \pm 0.02) \times 10^{-6}$ cm^2 s^{-1} at 0.1 and 3 mM concentration, respectively, indicating that the peptide does not aggregate at the higher concentration. Moreover, the 1D spectra of a 0.1 mM and of a 3.5 mM sample in H_2O were found to be practically identical. Therefore, we conclude that the $gp41_{659-671}$ peptide is monomeric in aqueous solution.

Circular Dichroism of $gp41_{659-671}$. The circular dichroism (CD) spectrum of $gp41_{659-671}$ was measured in water to assess the secondary structure of the peptide. The pH was adjusted to 7.7 by the addition of ammonium hydroxide. The

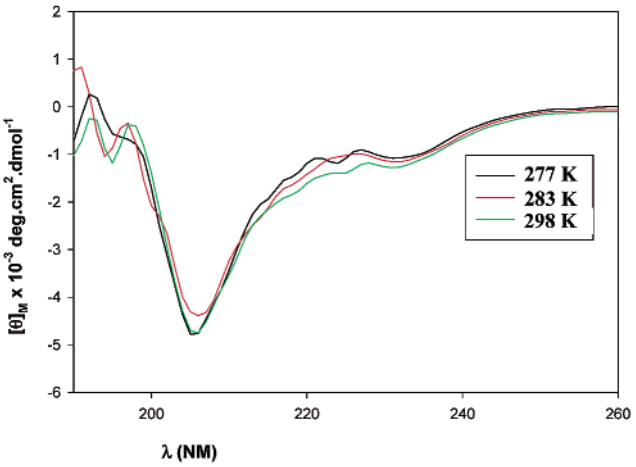

FIGURE 1: Circular dichroism of $gp41_{659-671}$ in H_2O. The CD spectra were measured at 277 K (black), 283 K (red), and 298 K (green). The peptide concentration was 0.4 mM in water, pH 7.7 at 277 K. CD intensities are expressed as mean residue ellipticities (deg·cm^2 dmol^{-1}).

CD curve was characterized by a weak negative shoulder between 220 and 230 nm, a minimum centered at 205 nm, very little ellipticity below 200 nm, and a ratio of $[\theta]_{222}/[\theta]_{205} = 0.27$ (Figure 1). Similar CD spectra were found for the peptide in ammonium acetate buffer at pH 7.7 and 8.4. The presence of two tryptophan residues in $gp41_{659-671}$ complicates the analysis, as the indole chromophore is known to contribute ellipticity in the peptide region of the CD spectrum (*33*). The shape and intensity of the CD patterns are similar to those recently reported for short peptides containing C^α-tetrasubstituted amino acids, which assume a 3_{10}-helical conformation (*24*, *34*). Moreover, a recently published CD spectrum of T-20 is very similar to the CD spectrum reported for the 3_{10}-helix (*35*). Overall, the CD results indicate that $gp41_{659-671}$ forms a 3_{10}-helix in aqueous media. No significant changes in the CD spectrum were observed at 283 and 298 K (Figure 1), indicating that the 3_{10}-helical conformation is mostly conserved over this range of temperatures.

NMR Analysis of $gp41_{659-671}$. $gp41_{659-671}$ was investigated by ^{1}H NMR in aqueous solution at pH 7.7 and 277 K, and chemical shifts were assigned by standard procedures. The NMR parameters are consistent with a helical structure for the segment ELDKWASLW (residues 662−670). Specifically, the peptide exhibited a series of $d_{NN(i,i+1)}$ and $d_{\alpha N(i,i+3)}$ NOE connectivities starting at E662 and continuing to the carboxyl end of the peptide (Figures 2a,b and Figure 3) and $^3J_{HN\alpha}$ values under 5.6 Hz with the exception of W670 and N671 that comprise the C-terminus of $gp41_{659-671}$ (Figure 3). The amide proton of L660 was not detected due to fast exchange with H_2O, frequently observed at this pH for the second amide proton in the polypeptide chain. In addition, the peptide shows negative deviations of α-proton chemical shifts from the random coil values (Figure 3). Both the $^3J_{HN\alpha}$ coupling constants, predicted to be 3.9 and 4.2 Hz, respectively (*25*), for ideal α- and 3_{10}-helices, and the H chemical shift deviations from random coil values support the presence of a high population of helical conformers in solution and only a minor population of random coil. The NOESY spectrum also exhibited 7 $d_{\alpha N(i,i+2)}$ connectivities but no $d_{\alpha N(i,i+4)}$ interactions (Figures 2a and 3). This pattern of

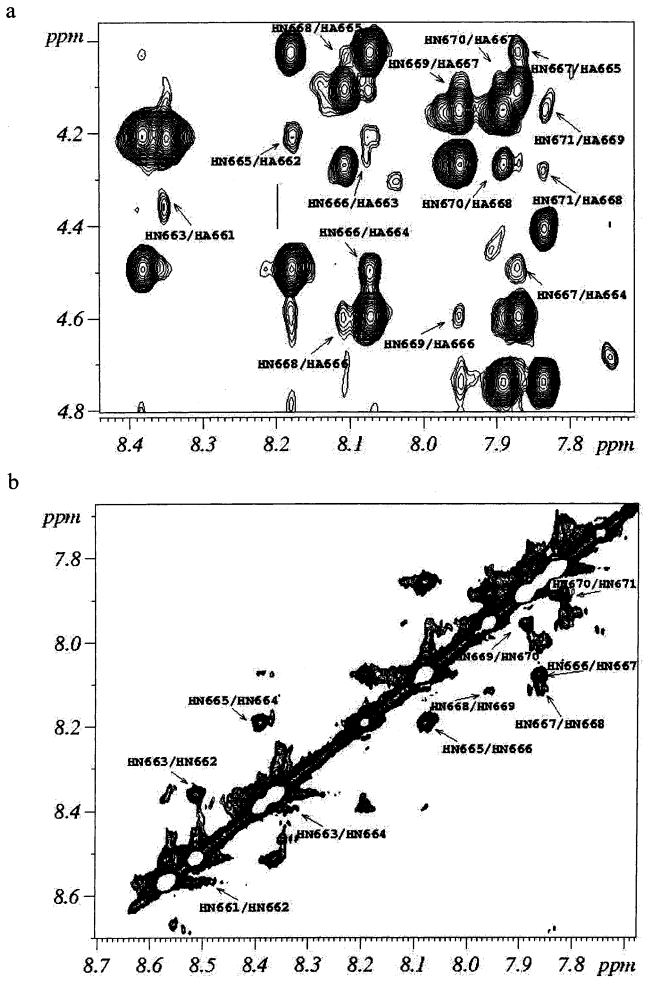

FIGURE 2: NMR NOESY spectrum of gp41$_{659-671}$. (a) The HN−H fingerprint region. (b) The HN−HN region. The concentration of gp41$_{659-671}$ was 1.85 mM in H_2O/D_2O (95:5), pH 7.7. The spectrum was recorded at 277 K with 250 ms mixing time on a Bruker DRX 800 MHz spectrometer. Cross-peaks are labeled using numbers indicating the position of the residue in gp160, the precursor for gp120, and gp41.

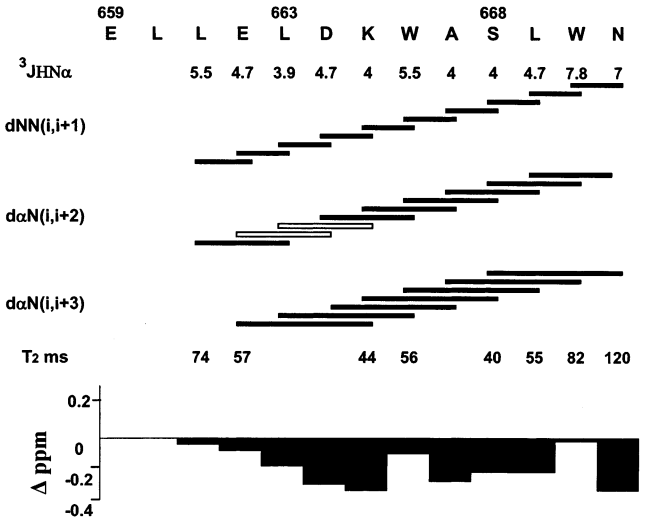

FIGURE 3: Analysis of NMR parameters of gp41$_{659-671}$. A summary of observed $^3J_{HN\alpha}$ coupling constants, interresidue connectivities, T_2 values in milliseconds, and deviation of the H chemical shifts from the random coil values. Observed interresidue connectivities are marked with filled rectangles, and the connectivities that could not be unambiguously assigned due to overlaps are marked with empty rectangles.

medium-range connectivities is indicative of *i*,*i*+3 hydrogen bonds and the presence of a 3_{10}-helix and rules out the presence of *i*,*i*+4 hydrogen bonds and an α-helical conformation in gp41$_{659-671}$ (*25*, *26*).

The stability of the helical conformation was studied by measuring the $^3J_{HN\alpha}$ coupling constants from 500 MHz 1D spectra recorded under different conditions. Both an increase in pH (from 7 to 8) and the addition of 50−100 mM Tris−acetate or ammonium acetate buffer at pH 7.7 caused a decrease in $^3J_{HN\alpha}$ values, indicative of greater helical stability. On the other hand, addition of 50 mM NaCl or elevation of the temperature resulted in a slight increase of the $^3J_{HN\alpha}$ coupling constants, suggesting somewhat lower helical stability. The T_2 relaxation time of the amide protons of residues 662−669 was found to be 40−57 ms while residues 661 and 670−671 exhibited T_2 in the range 74−120 ms, indicating a tighter conformation in the segment 662−669 and enhanced flexibility at the peptide termini (Figure 3).

Structure of the gp41 Peptide. The three-dimensional structure of the peptide was determined using 202 distance constraints and 9 ϕ and 4 χ_1 dihedral angle constraints. No hydrogen bond constraints were included initially. However, one *i*,*i*+3 hydrogen bond (between D664 and A667) was observed in all of the accepted structures (the created hydrogen bond was up to 2.8 Å in length with up to a 35° deviation from linearity), and the segment E662−W670 in the average energy-minimized structure was found to form a 3_{10}-helix conformation with 3.04 residues per turn. To gain experimental evidence for the presence of H-bonds in gp41$_{659-671}$, we carried out deuterium exchange experiments in D_2O. However, all protons exchanged at the earliest time point we could examine (approximately 1 h after dissolving the peptide in D_2O; data not shown). This is not surprising considering the high pH (7.8) of the sample and the increased solvent penetration in peptides in comparison to proteins. In addition, we measured the NH chemical shifts of the peptide in water over the temperature range from 277 to 297 K in 5 deg intervals. The data were linear, and we found a very low temperature coefficient (−2 ppb/K) for the A667 amide proton. Several other amides of gp41$_{659-671}$ had temperature coefficients ≤−5 ppb/K. These results are consistent with the presence of a strong hydrogen bond involving the A667 amide proton and several other H-bonded or solvent-shielded protons in the peptide. We note that in GGXGG peptides where X is an aromatic residue, in particular Trp, anomalously low NH temperature coefficients for the carboxyl-terminal Gly, compared to random coil values, were reported (*36*). However, the primary sequence of gp41$_{659-671}$ is not expected to lead to such a deviation for A667.

On the basis of the above calculation and the temperature coefficient investigation we input an *i*,*i*+3 hydrogen bond constraint between D664 and A667 into the final structure calculation. A superposition of the 25 lowest energy peptide structures, out of 100 calculated structures, and the average structure are presented in panels a and b of Figure 4, respectively. The structural statistics are presented in Table 1. The peptide exhibits a regular 3_{10}-helix conformation in the segment E662−W670 with 3.2 residues per turn, consistent with the structural features observed in a survey of 3_{10}-helices in protein crystal structures (*21*). The arrange-

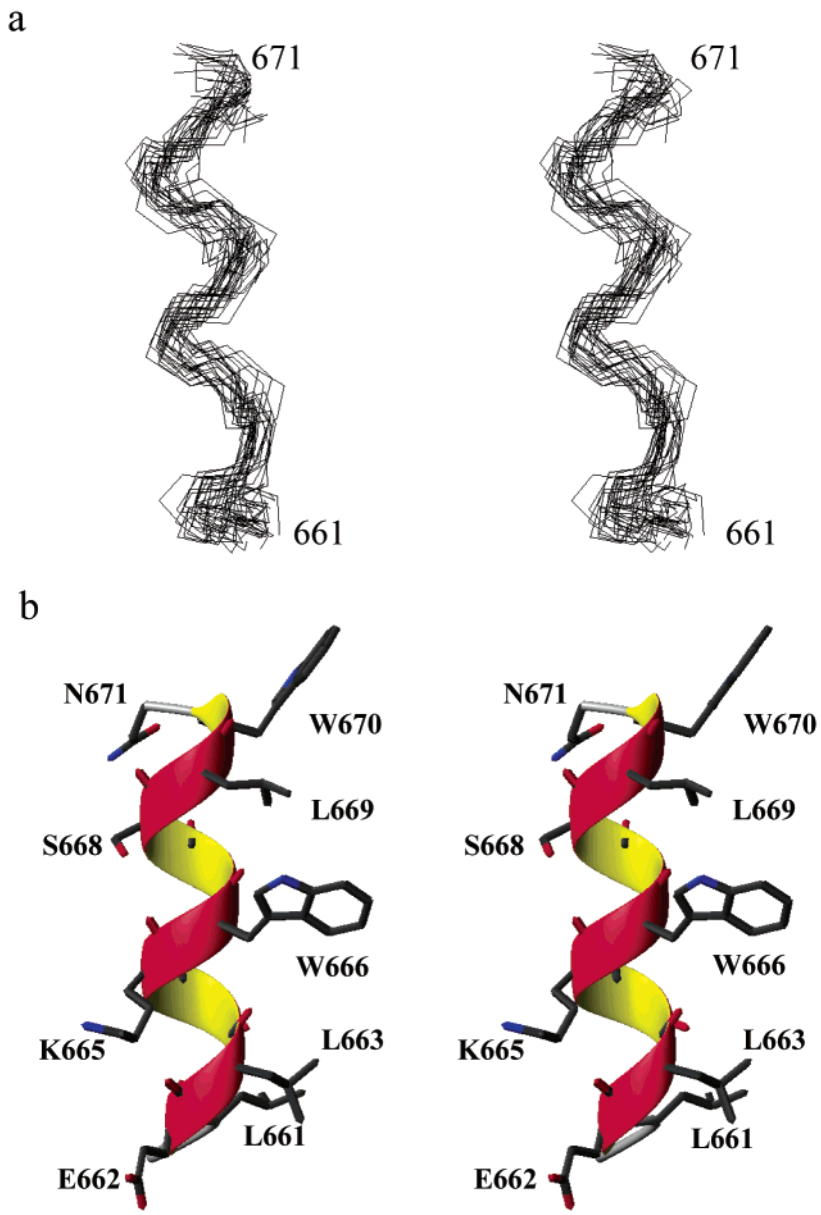

FIGURE 4: NMR-derived solution structure of $gp41_{659-671}$. (a) Superposition of 25 lowest energy structures calculated using NOE and dihedral angle constraints. (b) Stereoview of the average energy-minimized structure (average of 96 accepted structures).

ment of 3.2 residues per turn in the 3_{10}-helix formed by $gp41_{659-671}$ improves the staggering of the side chains in comparison to a 3_{10}-helix with 3 residues per turn (*21*). Two-thirds of the surface of the helix consisted of hydrophobic and tryptophan side chains. The conserved L663, W666, L669, and W670 residues (*8*) form a tightly packed hydrophobic patch perturbed somewhat by the indole of W666 and W670 as shown in Figure 5a. E662 and the variable K665, S668, and N671 (*8*) form a polar face (Figure 5a), and E662 and K665 may have weak electrostatic interactions between them. The role of this conserved surface is yet to be investigated. The helical wheels shown in Figure 5b,c demonstrate that the amphiphilic structure is formed by a 3_{10}-helix but not by an α-helix. The burial of considerable surface area of the hydrophobic side chains and of the tryptophans by the formation of the 3_{10}-helix explains the water solubility of $gp41_{659-671}$. An amphiphilic 3_{10}-helix could be characterized by the triad repeat NNP, where P and N are polar and nonpolar residues, respectively. The NNP motif in $gp41_{659-671}$ is broken by two tryptophan residues and an aspartic acid residue, which may explain why $gp41_{659-671}$ does not oligomerize. A steric hindrance caused by the bulky and rigid indoles of the two tryptophan residues may further impede oligomerization. E659, L660, and N671 did not exhibit any ordered secondary structure probably due to chain-end effects. The NMR parameters indicate that L661 is probably an incipient 3_{10}-helix and that it may be in equilibrium with disordered structures.

Spectroscopic and Structural Characterization of N-Acetylated, C-Amidated $gp41_{659-671}$ (Ac-$gp41_{659-671}$-NH_2). To abolish the destabilization of a helical conformation caused by the positive and negative charges at the N- and C-terminus, respectively, and to make the peptide more similar to the corresponding segment in the native protein, we synthesized the peptide Ac-$gp41_{659-671}$-NH_2 acetylated at the N-terminus and with an amide instead of a carboxyl group at the C-terminus. The diffusion coefficient of Ac-$gp41_{659-671}$-NH_2 at 0.1 mM was identical to that of $gp41_{659-671}$, indicating

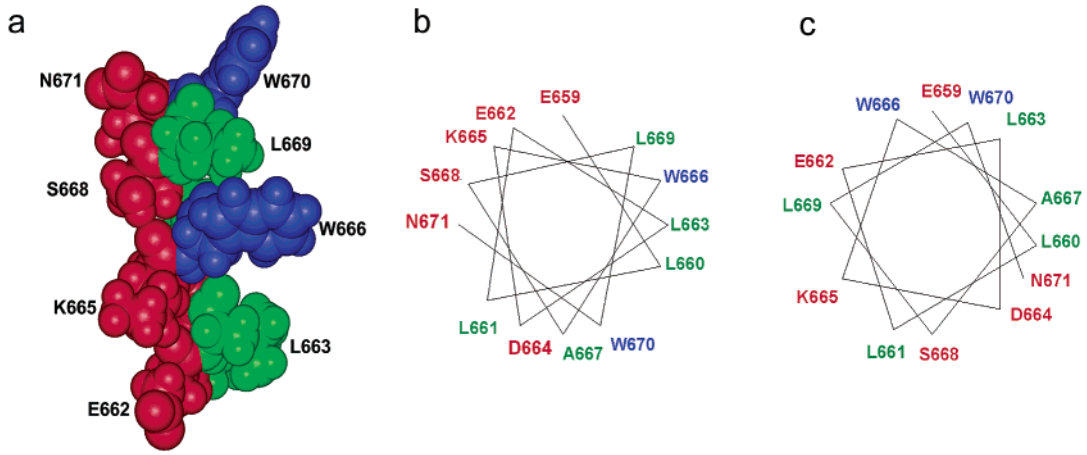

FIGURE 5: Amphiphilic structure of gp41$_{659-671}$. (a) NMR structure of gp41$_{659-671}$ shown in a CPK model. (b) A 3_{10}-helical wheel projection of gp41$_{659-671}$ with 3.2 residues per turn. (c) An α-helical wheel projection of gp41$_{659-671}$. Leucine, tryptophan, and polar residues are in green, blue, and red, respectively.

that Ac-gp41$_{659-671}$-NH_2 is monomeric at 0.1 mM concentration. At 3 mM the diffusion coefficient of Ac-gp41$_{659-671}$-NH_2 decreased by 7.5%, showing that the peptide remains mostly monomeric (dimerization reduces the diffusion coefficient by 20%). The CD spectrum exhibited a ratio $[\theta]_{222}/[\theta]_{205} = 0.37$ (data not shown), in the range (0.15−0.4) that is characteristic for 3_{10}-helices (*24*, *34*). The NMR data are consistent with the conclusion that the concurrent acetylation and amidation in Ac-gp41$_{659-671}$-NH_2 stabilize a helical conformation by eliminating charges at the termini that interacted unfavorably with the helical dipole. The amide protons of E659 and L660 were clearly observed in the spectrum of Ac-gp41$_{659-671}$-NH_2. As a result, an additional αN(*i*,*i*+2) connectivity between E659 and L661 could be observed (Figures 6a and 7), indicating a possible hydrogen bond between the terminal acetyl and the amide of L661, thus elongating the 3_{10}-helix. The helix stabilization manifests itself in lower $^3J_{HN\alpha}$ and stronger αN(*i*,*i*+3) NOEs in comparison to αN(*i*,*i*+2) NOEs (Figures 2a, 3, 6a, and 7). Numerous αN(*i*,*i*+2) NOEs and only one considerably weaker αN(*i*,*i*+4) NOE (between residues L663 and A667) were detected (see Figure 6a). Some very weak Nα(*i*,*i*+1) NOEs corresponding to a distance of 5.1 Å in a 3_{10}-helix were detected, suggesting that the αN(*i*,*i*+4) distances are longer than 5.1 Å. Altogether, these findings support the formation of *i*,*i*+3 rather than *i*,*i*+4 hydrogen bonds. The overlap between the amide protons of K665 and S668 and between W666 and W670 of Ac-gp41$_{659-671}$-NH_2 (see Figure 6a,b in comparison to Figure 2a,b) hampered the unambiguous assignment of several $d_{\alpha N(i,i+2)}$ and $d_{\alpha N(i,i+3)}$ NOEs. As a result the calculated structure of Ac-gp41$_{659-671}$-NH_2 was not as well defined as that of gp41$_{659-671}$. On the basis of the CD spectrum and the observation of pronounced αN(*i*,*i*+2) interactions and only a single and considerably weaker αN(*i*,*i*+4) NOE, we conclude that Ac-gp41$_{659-671}$-NH_2 forms mostly a 3_{10}-helix with a possible minor contribution of an α-helical conformation.

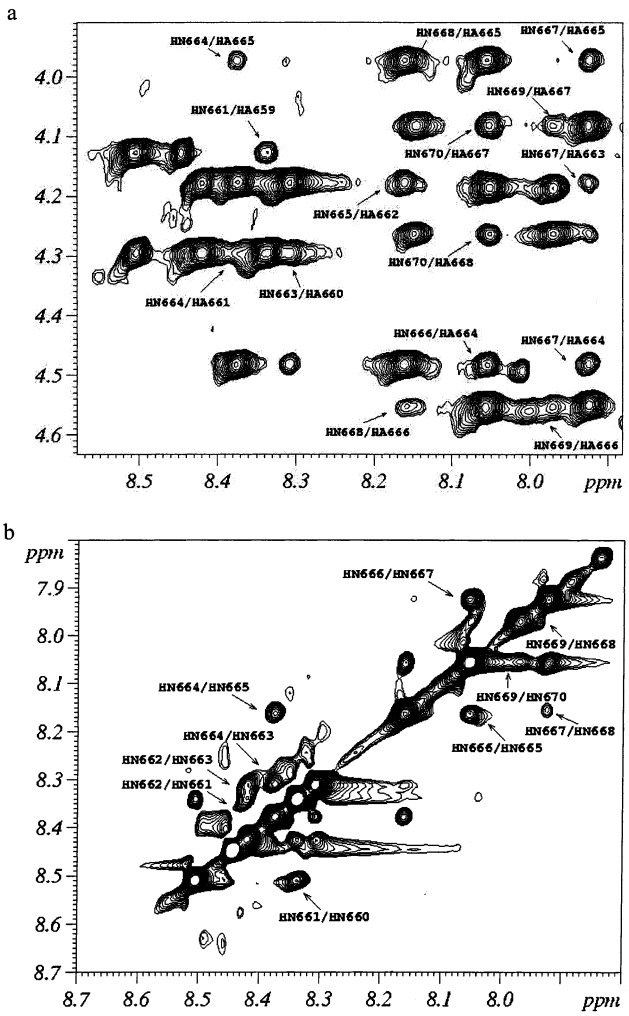

FIGURE 6: NMR NOESY spectrum of Ac-gp41$_{659-671}$-NH_2. (a) The HN−H fingerprint region. (b) The HN−HN region. The concentration of Ac-gp41$_{659-671}$-NH_2 was 1.85 mM in H_2O/D_2O (95:5), pH 7.7. The spectrum was recorded at 277 K with 250 ms mixing time on a Bruker DRX 800 MHz spectrometer. Cross-peaks are labeled using numbers indicating the position of the residue in gp160, the precursor for gp120, and gp41.

DISCUSSION

Comparison with Other Helical Peptides. It is very unusual for short peptides in aqueous solution to show considerable populations in helical conformations. Studies of α-helix formation by long polypeptides, together with the Zimm−Bragg model (*37*), predict that short protein fragments as well as other short peptides should not show any measurable helix formation in H_2O. The observations that a peptide

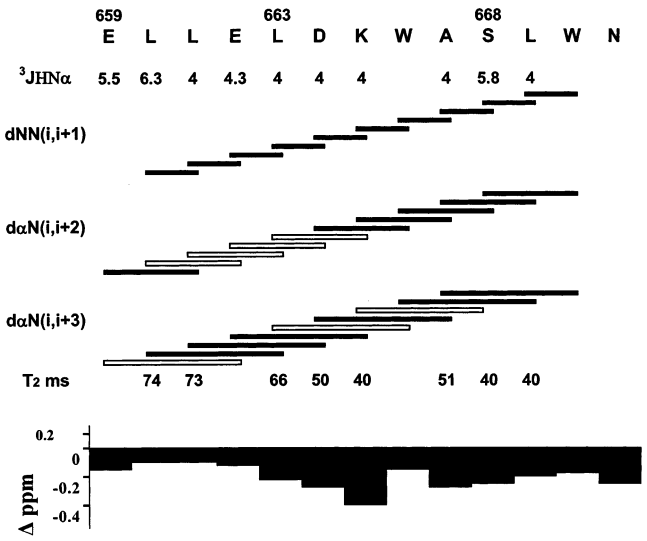

FIGURE 7: Analysis of NMR parameters of Ac-gp41$_{659-671}$-NH_2. A summary of observed $^3J_{HN\alpha}$ coupling constants, interresidue connectivities, T_2 values in milliseconds, and deviation of the H chemical shifts from the random coil values. Observed interresidue connectivities are marked with filled rectangles, and the connectivities that could not be unambiguously assigned due to overlaps are marked with empty rectangles.

comprising the first 13 residues of ribonuclease A, a poly(Lys-Ala-Ala) peptide, and a 17-mer peptide consisting of alanine and lysine residues form helices in water prompted numerous experimental and theoretical investigations of α-helix formation (*38*−*40*).

The 3_{10}-helix has long been thought to be inherently less stable than the α-helix since the packing of the backbone atoms is tighter and the hydrogen bonds are not parallel to the helix axis (*41*). As a result, the 3_{10}-helix is in most cases less than four residues long. The 3_{10}-helix has been primarily associated with peptides containing C^{α}-tetrasubstituted amino acids (*24*). Short α-helical peptides may partially adopt a 3_{10}-helix conformation especially at their termini (*26*). The peptide gp41$_{665-683}$, which partially overlaps gp41$_{659-671}$, is insoluble in water and is mostly α-helical in dodecylphosphocholine micelles (*42*). Some stretches in an ensemble of 40 calculated NMR structures for gp41$_{665-683}$ form either α- or 3_{10}-helices. The gp41$_{659-671}$ peptide studied herein is the first soluble peptide composed of natural amino acids found to form a stable 3_{10}-helical structure in water.

Interactions Stabilizing a 3_{10}- Rather than α-Helical Conformation in gp41$_{659-671}$. Favorable side chain interactions stabilize α-helices in aqueous solution (*43*−*47*). Theoretical calculations suggest that strong *i*,*i*+4 side chain interactions favor α-helix formation while the 3_{10}-helix population is maximized when weaker *i*,*i*+4 interactions are present (*48*). The AGADIR program (*49*) predicts only 0.58% and 2.57% α-helical content for gp41$_{659-671}$ and Ac-gp41$_{659-671}$-NH_2, respectively. In addition, examination of the amino acid sequences at the central section of long 3_{10}-helices (six residues and longer) shows that the propensity of different amino acids to form 3_{10}-helix decreases in the following order: Trp, Tyr, Asp, Glu, Met (*50*). Notably these propensities are quite different from those obtained for an α-helix.

Analysis of the gp41$_{659-671}$ sequence shows that the 3_{10}-helix rather than α-helix is expected to be favored, especially in the LLELDKW sequence, on the basis of the following considerations:

(a) A Leu-Leu pair at positions *i*,*i*+4 but not at positions *i*,*i*+3 stabilizes an α-helix as observed by CD spectroscopy (*46*). In gp41$_{659-671}$ L660 and L663 are at positions *i*,*i*+3.

(b) Theoretical calculations show that interactions involving three leucine residues at positions *i*,*i*+1 and *i*+4 are favorable for an α-helix (*47*). In gp41$_{659-671}$ the leucine triplet is *i*,*i*+1,*i*+3.

(c) The observed destabilization of the 3_{10}-helix at a higher salt concentration is indicative of a stabilizing Glu-Lys *i*,*i*+3 interaction in gp41$_{659-671}$. In an α-helix the Glu-Lys pair is stabilizing at positions *i*,*i*+4 but not at *i*,*i*+3 (*44*).

(d) The Leu-Trp pair stabilizes an α-helix in *i*,*i*+4 positions and destabilizes an α-helix when it is found in positions *i*,*i*+3, as in gp41$_{659-671}$ (*47*).

(e) Asp and Trp have low and medium propensities for an α-helix, respectively, and very high propensities for a 3_{10}-helix (*50*).

A leucine-rich repeat variant provides an excellent example in which favorable interactions caused by increasing the space between two leucines to two residues led to the formation of a 3_{10}-helix (*51*). The favorable *i*,*i*+3 interactions and the scarcity of favorable *i*,*i*+4 interactions as well as the significant number of amino acids with a high propensity for the 3_{10}-helix may have contributed to the unusual stabilization of the 3_{10}-helix conformation in gp41$_{659-671}$.

gp41$_{659-671}$ as a 3_{10}-Helix-Forming Model Peptide. The finding that the ribonuclease C−peptide forms an α-helix (*38*) was the impetus to a large number of studies seeking to understand the factors that influence α-helix formation in water. Similarly, the discovery that the gp41$_{659-671}$ and Ac-gp41$_{659-671}$-NH_2 peptides form a 3_{10}-helix can be exploited to conduct detailed studies of 3_{10}-helix formation and stability. Previous peptides that formed 3_{10}-helices were comprised of C^{α}-substituted amino acids, preventing the use of NOEs involving α-protons to differentiate between α-helical and 3_{10}-helical conformations (*23*). NOEs involving Cα protons can be easily measured for gp41$_{659-671}$, which is comprised only of amino acids found in proteins. Factors influencing the stability of the 3_{10}-helix such as the contribution of salt bridges and side chain interactions can thereby be studied in detail. The propensity of the different amino acids to form a 3_{10}-helix is not fully known and seems to differ from the propensity to form an α-helical conformation. Monitoring the effects of systematic variations of the gp41$_{659-671}$ sequence upon the 3_{10}-helix stability can help in evaluating the propensities of different amino acids to form a helical structure. Factors that influence the self-association of 3_{10}-helical peptides can be studied as well.

gp41$_{659-671}$ Is an Autonomous Folding Unit, and Its Structure Represents That of the Corresponding Segment in Native gp41. The observation that a peptide corresponding to a fragment of a protein has a very significant population in a helical conformation makes it extremely likely that the corresponding region in the cognate protein adopts the same conformation (*52*−*54*). A helix that is stable enough to be observed in an isolated peptide is an autonomous folding unit and may serve as a nucleation site during protein folding (*55*−*57*).

A heptapeptide (ELDKWAS, gp41$_{662-668}$) was found to adopt a *β*-turn formed by the DKWA sequence when bound

to the antibody 2F5 (*58*). Both the 3_{10}-helix and the β-turn are characterized by an *i*,*i*+3 hydrogen bond, unlike the α-helix which is characterized by *i*,*i*+4 hydrogen bonds. This structural similarity supports the hypothesis that the corresponding region of the native gp41 adopts a 3_{10}-helix conformation rather than an α-helix.

The backbone rmsd between the NMR structure of $gp41_{659-671}$ and the crystal structure of ELDKWAS ($gp41_{662-668}$) bound to 2F5 is 0.861 Å for the DKWA sequence and 2.619 Å for the seven-residue segment, indicating that the first two residues and the last residue of the antibody-bound peptide differ considerably in their conformation from that found in $gp41_{659-671}$. Furthermore, all of the NOE data we measured for the DKWA part of $gp41_{659-671}$ in water were consistent with the X-ray structure of this region of the 2F5 bound peptide. When we overlaid the X-ray structure onto our ensemble of $gp41_{659-671}$ structures, we were unable to accommodate the distorted type I β-turn of the crystal structure in our calculated structures. Specifically, there was a serious mismatch in the ϕ angle of the W666 residue in the NMR and X-ray structures.

Since the antibody 2F5 was elicited against a native virus, most likely it recognizes a conformation similar to the native conformation of gp41. However, several factors seriously undermine the ability of the heptapeptide to mimic the native gp41 conformation when bound to the antibody. Most of the interactions of the ELDKWAS peptide with the antibody are contributed by the sequence DKW. As a result, the first two and last residues of the heptapeptide may be only partially constrained by the interaction with the antibody. Moreover, cyclic peptides, in which the side chains of the second (L663) and sixth (A667) residues were modified and cross-linked via intramolecular disulfide or amide bond formation, could be crystallized in complex with 2F5 (*58*). The modification of the side chains of the second and sixth residue indicated that they are not crucial for binding, and therefore, it is questionable whether in the linear peptide they are constrained to the native gp41 conformation when bound to 2F5. The positively charged N-terminus and the negatively charged C-terminus do not mimic peptide bonds at these positions of gp41, and furthermore, these charges are expected to destabilize a helical conformation due to repulsion with the natural helix dipole.

The peptide conformation in the antibody complex is stabilized not only by the interactions with the antibody but also by the medium- and long-range intramolecular interactions within the bound peptide. As stated above, the stabilization of the 3_{10}-helix conformation requires favorable *i*,*i*+3 side chain interactions. Due to its short size many interactions that would stabilize a 3_{10}-helix in the corresponding seven residues in the longer 13-mer peptide cannot form in the heptapeptide. For example, while the *i*,*i*+3 interaction between L663 and W666 is possible in the heptapeptide, the stabilizing *i*,*i*+2 and *i*,*i*+3 interactions of L663 with L660 and L661 cannot occur as the latter two are absent in the heptapeptide. Similarly, the 3_{10}-helix stabilizing *i*,*i*+3 interactions of W666 with L669 cannot occur in the heptapeptide.

If the structure of a protein-derived peptide in complex with an antibody, a toxin, or a receptor protein is used to learn about the conformation of the corresponding segment in the native protein, it is of utmost importance to include not only the residues directly involved in intermolecular interactions but also the residues involved in medium- and long-range intramolecular interactions that define the secondary structure of the protein segment. We have demonstrated that mapping of the minimal determinant that can faithfully mimic the conformation of the corresponding part of the protein can be done by NMR dynamic filtering experiments (*59*). Using this approach we determined a β-hairpin conformation in an acetylcholine receptor peptide bound to α-bungarotoxin (*60*, *61*) while others who used shorter peptide fragments of the receptor did not find this bound conformation (*62*−*64*). In summary, the 13-residue $gp41_{659-671}$ peptide but not the gp41 heptapeptide contains the complete epitope recognized by 2F5 (*13*, *14*). In view of the above considerations we conclude that the heptapeptide lacks important interactions with the antibody as well as helix-stabilizing interactions within the peptide and therefore cannot mimic accurately the corresponding region of gp41. It is likely, however, that the NMR structure of $gp41_{659-671}$ closely resembles that of the same region in the coat protein.

The Structure of $gp41_{659-671}$ May Correspond to All Forms of gp41. The 2F5 antibody binds gp41 only in the absence of CD4 binding to the envelope protein. The interaction of the gp120/gp41 complex with the CD4 surface molecules of T-cells and macrophages leads to the exposure of other gp41 epitopes but apparently disrupts the conformation of the 2F5 epitope or reduces its exposure (*19*, *20*). Previous crystallographic studies determined that the structure of the gp41 core is a six-helix bundle conformation in which the N-terminal domain interacts with the C-terminal domain to form a trimer of helical hairpins. This conformation, induced after CD4 binding and the dissociation of the gp120/gp41 complex, most likely represents the fusion-active conformation of gp41. The unusual stability of the 3_{10}-helix conformation in the isolated peptide may indicate that the helical conformation is maintained in gp41 throughout the different conformations of this envelope protein. However, the exposure of the $gp41_{659-671}$ segment may vary in the different gp41 conformations, leading to changes in the reactivity with 2F5.

Contrary to the assumed flexibility of the 18-residue link ($gp41_{666-683}$) between the C-peptide region ($gp41_{628-665}$) and the transmembrane anchor in the fusogenic or postfusion form of gp41 (*2*), we show that at least six residues of this link adopt a helical structure. The absence of a strong self-association between $gp41_{659-671}$ strands may allow the interdomain flexibility required for the formation of the hairpin structure of the gp41-core. This flexibility would permit the packing of the C-segment against the surface of the coiled coil formed by the N-segment as observed in the crystal structure (*1*−*3*). Additional studies on peptides corresponding to larger gp41 segments should provide new information that will further illuminate the fusion process.

Implications for the Development of an Anti-HIV-1 Vaccine. The peptide $gp41_{659-671}$ represents the complete epitope for the antibody 2F5 (*13*, *14*), which is one of the only three antibodies capable of efficient, broad-spectrum neutralization of HIV-1 infection. 2F5 is the only one of the three antibodies (2F5, 2G12, and G1b12) (*15*) that recognizes a sequential determinant. Peptide immunogens have the potential to be used as a component of a synthetic vaccine against HIV-1. It has been proposed that the best peptide immunogens for

eliciting antibodies cross-reactive with native proteins are those peptides that have a high folding propensity (*52*, *65*). The probability to induce protein-reactive antibodies increases with the length of the peptide immunogen (*66*), and a minimum of 10 residues may be required (*52*). Induction of protein-reactive antibodies can be optimized by adjusting the peptide length to that required for stabilizing a nativelike conformation of the peptide. The length of such peptides may differ according to the secondary structure (β-turn, helix, or β-hairpin) and according to the propensity of the peptide to form this conformation. Autonomous folding units fulfill this requirement while shorter peptides will not be as effective in eliciting protein-reactive antibodies.

Indeed, attempts to elicit efficient HIV-1 neutralizing antibodies with the ELDKWA sequence in various forms such as insertion into recombinant proteins or in multiple antigen peptides have failed (*13*, *67*, *68*). This failure led to the hypothesis that the conformation of the epitope is dependent on the environment in which it is present in gp41 or that ELDKWA is not the complete epitope (*13*). To overcome this problem, the use of constrained peptides has been suggested (*69*, *70*). The finding that $gp41_{659-671}$ is an autonomous folding unit may have profound implications on HIV-1 vaccine development, since peptide immunogens consisting of the complete $gp41_{659-671}$ sequence will adopt a 3_{10}-helical conformation without any artificial structural constraints and are, therefore, likely to induce antibodies that are highly cross-reactive with HIV-1. Moreover, unstructured C-peptide immunogens, which include the $gp41_{659-671}$ sequence, may not elicit broadly neutralizing antibodies because their linear sequence contains some residues that are variable among different HIV-1 strains (*69*). The folding of $gp41_{659-671}$ as a monomeric 3_{10}-helix enables the exposure of the mostly hydrophobic surface to the immune system while the hydrophilic variable residues form the opposite face. This mostly hydrophobic surface contains the conserved tryptophan and leucine residues that can potentially elicit antibodies that cross-react with a broad spectrum of HIV-1 isolates.

Our results ratify strategies for development of synthetic anti-HIV vaccines directed against antigenic determinants with varying exposure during natural infection. At the same time it must be noted that both immunization with the $gp41_{659-671}$ epitope and passive immunization with the 2F5 antibody are very likely to interfere with T-20 treatment and, thus, may compromise a promising route to the next generation of anti-HIV-1 drugs.

ACKNOWLEDGMENT

We are very grateful to Mr. Rotem Sorek for conducting the computer search of the SwissProt, to Dr. Deborah Fass for help in the sedimentation experiments, to Drs. Leonid Konstantinovskii and Yoram Cohen for help in the diffusion measurements, to Jordan Chill and Drs. Deborah Fass and Joel Sussman for helpful comments, and to Dr. Jana Wehrle for suggesting this project.

REFERENCES

1. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell 89*, 263−273.
2. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature 387*, 426−430.
3. Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 12303−12308.
4. 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.
5. Chen, J., Skehel, J. J., and Wiley, D. C. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 8967−8972.
6. Chan, D. C., and Kim, P. S. (1998) *Cell 93*, 681−684.
7. Nabel, G. J. (2001) *Nature 410*, 1002−1007.
8. Salzwedel, K., West, J. T., and Hunter, E. (1999) *J. Virol. 73*, 2469−2480.
9. Poumbourios, P., el Ahmar, W., McPhee, D. A., and Kemp, B. E. (1995) *J. Virol. 69*, 1209−1218.
10. Lawless, M. K., Barney, S., Guthrie, K. I., Bucy, T. B., Petteway, S. R., Jr., and Merutka, G. (1996) *Biochemistry 35*, 13697−13708.
11. Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, L., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M., and Saag, M. S. (1998) *Nat. Med. 4*, 1302−1307.
12. Purtscher, M., Trkola, A., Gruber, G., Buchacher, A., Predl, R., Steindl, F., Tauer, C., Berger, R., Barrett, N., Jungbauer, A., et al. (1994) *AIDS Res. Hum. Retroviruses 10*, 1651−1658.
13. Parker, C. E., Deterding, L. J., Hager-Braun, C., Binley, J. M., Schulke, N., Katinger, H., Moore, J. P., and Tomer, K. B. (2001) *J. Virol. 75*, 10906−10911.
14. Conley, A. J., Kessler, J. A. N., Boots, L. J., Tung, J. S., Arnold, B. A., Keller, P. M., Shaw, A. R., and Emini, E. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3348−3352.
15. Burton, D. R., and Montefiori, D. C. (1997) *AIDS 11*, S87−S98.
16. Mascola, J. R., Stiegler, G., VanCott, T. C., Katinger, H., Carpenter, C. B., Hanson, C. E., Beary, H., Hayes, D., Frankel, S. S., Birx, D. L., and Lewis, M. G. (2000) *Nat. Med. 6*, 207−210.
17. Baba, T. W., Liska, V., Hofmann-Lehmann, R., Vlasak, J., Xu, W. D., Ayehunie, S., Cavacini, L. A., Posner, M. R., Katinger, H., Stiegler, G., Bernacky, B. J., Rizvi, T. A., Schmidt, R., Hill, L. R., Keeling, M. E., Lu, Y. C., Wright, J. E., Chou, T. C., and Ruprecht, R. M. (2000) *Nat. Med. 6*, 200−206.
18. Burton, D. R., and Parren, P. W. (2000) *Nat. Med. 6*, 123−125.
19. Sattentau, Q. J., Zolla Pazner, S., and Poignard, P. (1995) *Virology 206*, 713−717.
20. Gorny, M. K., and Zolla-Pazner, S. (2000) *J. Virol. 74*, 6186−6192.
21. Barlow, D. J., and Thornton, J. M. (1988) *J. Mol. Biol. 201*, 601−619.
22. Millhauser, G. L. (1995) *Biochemistry 34*, 3873−3877.
23. Dehner, A., Planker, E., Gemmecker, G., Broxterman, Q. B., Bisson, W., Formaggio, F., Crisma, M., Toniolo, C., and Kessler, H. (2001) *J. Am. Chem. Soc. 123*, 6678−6686.
24. Formaggio, F., Crisma, M., Rossi, P., Scrimin, P., Kaptein, B., Broxterman, Q. B., Kamphuis, J., and Toniolo, C. (2000) *Chemistry 6*, 4498−4504.
25. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
26. Millhauser, G. L., Stenland, C. J., Hanson, P., Bolin, K. A., and van de Ven, F. J. (1997) *J. Mol. Biol. 267*, 963−974.
27. Wu, D. H., Chen, A. D., and Johnson, C. S. (1995) *J. Magn. Reson., Ser. A 115*, 260−264.
28. Macura, S., and Ernst, R. R. (2002) *Mol. Phys. 100*, 135−147.
29. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
30. Roberts, G. C. K. (1993) *NMR of Macromolecules, a Practical Approach*, Oxford University Press, New York.
31. 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., Sect. D 54*, 905−921.
32. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 239*, 129−136.
33. Grishina, I. B., and Woody, R. W. (1994) *Faraday Discuss.*, 245−262.
34. Toniolo, C., Polese, A., Formaggio, F., Crisma, M., and Kamphuis, J. (1996) *J. Am. Chem. Soc. 118*, 2744−2745.
35. Mobley, P. W., Pilpa, R., Brown, C., Waring, A. J., and Gordon, L. M. (2001) *AIDS Res. Hum. Retroviruses 17*, 311−327.

36. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.
37. Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I. M., Stewart, J. M., and Baldwin, R. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2349−2353.
38. Brown, J. E., and Klee, W. A. (1971) *Biochemistry 10*, 470−476.
39. Yaron, A., Tal, N., and Berger, A. (1972) *Biopolymers 11*, 2461−2481.
40. Marqusee, S., Robbins, V. H., and Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286−5290.
41. Creighton, T. E. (1995) *Proteins, Structures and Molecular Properties*, 2nd ed., W. H. Freeman and Co., New York.
42. Schibli, D. J., Montelaro, R. C., and Vogel, H. J. (2001) *Biochemistry 40*, 9570−9578.
43. Lotan, N., Yaron, A., and Berger, A. (1966) *Biopolymers 4*, 365−368.
44. Marqusee, S., and Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898−8902.
45. Padmanabhan, S., and Baldwin, R. L. (1994) *J. Mol. Biol. 241*, 706−713.
46. Padmanabhan, S., and Baldwin, R. L. (1994) *Protein Sci. 3*, 1992−1997.
47. Creamer, T. P., and Rose, G. D. (1995) *Protein Sci. 4*, 1305−1314.
48. Sun, J. K., and Doig, A. J. (1998) *Protein Sci. 7*, 2374−2383.
49. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) *J. Mol. Biol. 284*, 173−191.
50. Pal, L., and Basu, G. (1999) *Protein Eng. 12*, 811−814.
51. Peters, J. W., Stowell, M. H., and Rees, D. C. (1996) *Nat. Struct. Biol. 3*, 991−994.
52. Dyson, H. J., Lerner, R. A., and Wright, P. E. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 305−324.
53. Munoz, V., and Serrano, L. (1996) *Folding Des. 1*, R71−R77.
54. Serrano, L. (2000) *Adv. Protein Chem. 53*, 49−85.
55. Chakrabartty, A., and Baldwin, R. L. (1995) *Adv. Protein Chem. 46*, 141−176.
56. Shoemaker, K. R., Fairman, R., Kim, P. S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 391−398.
57. Baldwin, R. L., and Rose, G. D. (1999) *Trends Biochem. Sci. 24*, 26−33.
58. Pai, E. F., Klein, M. H., Chong, P., and Pedyczak, A. (2000) in World Intellectual Property Organization (www.wipo.org) Patent WO-00/61618.
59. Zvi, A., Kustanovich, I., Feigelson, D., Levy, R., Eisenstein, M., Matsushita, S., Richalet Secordel, P., Regenmortel, M. H., and Anglister, J. (1995) *Eur. J. Biochem. 229*, 178−187.
60. Samson, A. O., Chill, J. H., Rodriguez, E., Scherf, T., and Anglister, J. (2001) *Biochemistry 40*, 5464−5473.
61. Samson, A. O., Scherf, T., Eisenstein, M., Chill, J. H., and Anglister, J. (2002) *Neuron 35*, 319−332.
62. Basus, V. J., Song, G., and Hawrot, E. (1993) *Biochemistry 32*, 12290−12298.
63. Zeng, H., Moise, L., Grant, M. A., and Hawrot, E. (2001) *J. Biol. Chem. 276*, 18.
64. Moise, L., Piserchio, A., Basus, V. J., and Hawrot, E. (2002) *J. Biol. Chem. 277*, 12406−12417.
65. Dyson, H. J., and Wright, P. E. (1995) *FASEB J. 9*, 37−42.
66. Tanaka, T., Slamon, D. J., and Cline, M. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3400−3404.
67. Coeffier, E., Clement, J. M., Cussac, V., Khodaei-Boorane, N., Jehanno, M., Rojas, M., Dridi, A., Latour, M., El Habib, R., Barre-Sinoussi, F., Hofnung, M., and Leclerc, C. (2000) *Vaccine 19*, 684−693.
68. Ho, J., MacDonald, K. S., and Barber, B. H. (2002) *Vaccine 20*, 1169−1180.
69. Root, M. J., Kay, M. S., and Kim, P. S. (2001) *Science 291*, 884−888.
70. Cabezas, E., Wang, M., Parren, P. W., Stanfield, R. L., and Satterthwait, A. C. (2000) *Biochemistry 39*, 14377−14391.

BI026261Y