HIV-1 gp41 Tertiary Structure Studied by EPR Spectroscopy†

Mark David Rabenstein and Yeon-Kyun Shin*

Department of Chemistry, University of California, Berkeley, California 94720

Received July 16, 1996⊗

ABSTRACT: HIV gp41 is the transmembrane glycoprotein responsible for fusion of viral and cellular membranes, enabling viral entry. The structure of gp41 was studied using two synthetic peptides derived from the ectodomain of gp41: a 38-residue peptide from the "heptad repeat" region (hr.wt), and a 34-residue peptide from a region closer to the C-terminus (bt.wt). These peptides were found to form a trimer of heterodimers with approximately 80% α-helicity. To study their alignment, distances between spin-labels attached to Cys residues on Cys-substituted peptides were measured using a recently-developed electron paramagnetic resonance method [Rabenstein, M. D., & Shin, Y.-K. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92,* 8239−8243]. The heterotrimeric peptides were found to be antiparallel, consistent with a study on proteolytically cleaved peptide fragments of gp41 [Lu, M., Blacklow, S. C., & Kim, P. S. (1995) *Nat. Struct. Biol. 2,* 1075−1082]. Furthermore, the C-terminal 19 residues of hr.wt are not apposed to bt.wt, and 15 residues of bt.wt extend beyond the end of hr.wt. Consistent with this alignment are tertiary interactions between specific sites of these peptides probed by spin-label mobility. Additionally, a second pair of peptides was studied. From the model, these are expected to align with complete overlap. Alone, neither was helical, but when mixed, they were 83% helical. Based on the alignment of the peptides, a model of the prefusogenic form of gp41 was constructed which is significantly different from the structure of influenza hemagglutinin.

On the surface of human immunodeficiency virus type 1 (HIV-1)[1] are the envelope (*env*) glycoproteins gp120 and gp41. gp120 and gp41 are synthesized as a single precursor, gp160, which is subsequently cleaved by a cellular protease. After proteolytic cleavage, they remain noncovalently associated. gp120 is the surface attachment subunit (SU), which recognizes the human CD4 receptor, and gp41 is the transmembrane domain (TM), which contains an N-terminal hydrophobic fusion peptide and is responsible for viral−cell membrane fusion. The structure of neither protein is known.

A "leucine zipper" motif, which is often responsible for oligomerization of proteins via coiled-coil formation, was discovered in the N-terminal region of the gp41 sequence (Gallaher *et al.*, 1989; Delwart *et al.*, 1990). It was postulated that this region is responsible for oligomerization of gp41 via coiled-coil formation (Gallaher *et al.*, 1989; Delwart *et al.*, 1990). "Leucine zippers" are characterized by a 3−4 heptad repeat of predominantly Leu, Ile, and Val residues. The residues are labeled from "a" to "g", with the "a" and "d" residues being predominantly Leu, Ile, and Val (Parry, 1982). A synthetic 38-residue peptide (dp107, or hr, see Figure 1) derived from the leucine zipper region of gp41 forms a tetrameric coiled-coil (Wild *et al.*, 1992; Rabenstein & Shin, 1995a). However, Lu *et al.* (1995) found that two peptides from gp41, one of which contains the heptad repeat region, form a trimer of heterodimers.

Here we report data on the tertiary structure of gp41. We studied the interaction between hr and a 34-residue peptide from near the transmembrane region (bt, see Figure 1). When mixed 1:1, these peptides were found to form a trimer of heterodimers with approximately 80% α-helicity. hr presumably forms a central coiled-coil, with bt packing against it as an α-helix (Lu *et al.*, 1995). We determined the position of bt relative to hr by measuring the distances between spin-labels on hr and bt peptides using a method described previously (Rabenstein & Shin, 1995b). This is the first reported application of this method. According to the distances measured, the peptides are antiparallel with 15 residues of bt extending beyond the end of hr. Tertiary contacts between the two peptides, inferred from the mobility of spin-labels at various locations of hr, were consistent with this alignment. A second pair of peptides, bt and hr2 (Figure 1), was expected to align with complete overlap. Alone, neither was helical. However, the two peptides formed an 83% α-helical complex. Based on this alignment and on previous studies, a model of the prefusogenic form of gp41 is presented, and differences between gp41 and influenza hemagglutinin (HA) are discussed.

MATERIALS AND METHODS

Peptides. Peptides were synthesized, purified, and spin-labeled with 3-(2-iodoacetamido)-PROXYL as described previously (Rabenstein & Shin, 1995a).

CD Spectroscopy. CD spectra were measured at 22 °C in 150 mM sodium chloride, 50 mM potassium phosphate, pH 7.2, and 10 μM peptide (unless otherwise noted) using a Jasco J600 CD spectrometer and a 5.0 mm path length cell. Peptide concentrations were determined by ultraviolet ab-

† Supported by the startup fund from the University of California at Berkeley, by NIH Grant GM51290-01, and by a University of California President's graduate student grant for research. M.D.R. is a recipient of an NIH training grant, and Y.-K.S. is a 1995 Searle Scholar.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1996.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; gp, glycoprotein; EPR, electron paramagnetic resonance; SU, surface attachment subunit; TM, transmembrane subunit; HA, influenza virus hemagglutinin; HPLC, high-performance liquid chromatography; MOPS, 3-(*N*-morpholino)propanesulfonic acid; hr, heptad repeat; wt, wild type; bt, buttressing; CD, circular dichroism; τ_c, rotational correlation time.

S0006-2960(96)01743-6 CCC: $12.00

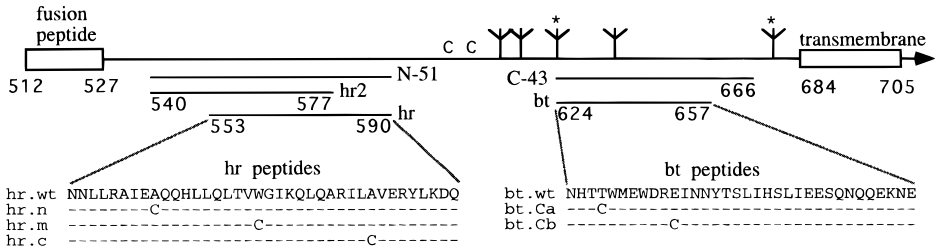

FIGURE 1: Diagram of HIV-1_{HXB2R} gp41 sequence, and peptides used in this study. The single-letter amino acid code is used to denote amino acid residues. hr and bt peptides are acetylated and amidated. Nitroxide spin-labels were attached to the Cys residues. The locations of N-51 and C-43 (Lu *et al.*, 1995) are also indicated. The conserved Cys residues are marked by a "C", and conserved glycosylation sites are marked. Nonconsensus sites are indicated by an asterisk.

sorption either at 280 or at 275.5 nm (Edelhoch, 1967). [θ_{222}] values are in units of degrees centimeter squared per decimole.

Analytical Ultracentrifugation. Sedimentation equilibria were determined with a Beckman Optima XL-A ultracentrifuge at 30K rpm, with a radius from 5.80 to 7.25 cm. The peptides were in 150 mM sodium chloride, 50 mM potassium phosphate, pH 7.2 at 22 °C. The solvent density and the partial specific volumes were calculated by standard methods (Weast, 1975; McMeekin *et al.*, 1949).

EPR Spectroscopy. EPR measurements were performed using a Bruker ESP300 EPR spectrometer (Bruker, Germany) equipped with a low-noise microwave amplifier (Miteq) and a loop−gap resonator (Medical Advances). The modulation amplitude was set at no greater than one-fifth of the line width. hr and bt peptides were mixed 1:1 with no buffer present. Buffer was added, and the samples were concentrated to 70−200 μM using a Microcon (Amicon) microconcentrator with a 10 kDa cutoff. EPR spectra were independent of concentration in this range. To control the temperature, the loop−gap resonator was inserted into a home-made quartz vacuum Dewar and cooled with nitrogen controlled by a Eurotherm B-VT 2000 thermostat.

Spin-Label Separation Calculations. The EPR method used to measure distances between two nitroxides was reported previously (Rabenstein & Shin, 1995b). Briefly, in the frozen state, a pair of nitroxides separated by the distance r are coupled to each other by electron−electron dipolar interactions. This causes a splitting dependent on their separation and their orientation with respect to the magnetic field:

$$\text{splitting} = (3/2)g_e\beta(3\cos^2\theta - 1)/r^3 \qquad (1)$$

where θ is the angle between the interspin vector and the magnetic field, g_e the electron g value, and β the Bohr magneton. Because the sample is noncrystalline, θ is isotropically distributed, leading to broadening instead of splitting of the spectrum. Furthermore, it is possible to extract the broadening function by Fourier deconvolution of the broadened and unbroadened EPR spectra. The unbroadened EPR spectrum, or noninteracting spectrum, is the spectrum of a spin-label on a monomeric peptide or protein. Using α-helical model peptides, this method was found to have 1 Å accuracy within the range 8−25 Å (Rabenstein & Shin, 1995b).

The measurement technique was developed for measuring the distance between two spin-labels on a monomeric protein. However, the hr/bt complex (see the first paragraph of the Results section for nomenclature) is a 6-spin system, as there are spin-labels on each of the peptides in the trimer of heterodimers. We wanted to measure specifically the distance between a spin-label on an hr peptide and a spin-label on a bt peptide. Interactions between spin-labels on peptides of the same kind (i.e., two spin-labels on different hr peptides within the same complex) can also broaden the EPR spectrum. For example, the spectrum of hr.m/bt.wt is broader than that of a spin-label on a monomeric peptide, showing that spin-labels on hr.m peptides within the complex interact with each other (Figure 4). However, because the broadening is dependent on the reciprocal of the cube of the separation (eq 1), spin-label pairs with short separations have a much greater effect on the spectrum than those with longer separations.

As a first approximation, we assumed that the broadening from different spins in a multispin system is additive. By independently estimating the broadening due to interactions between spin-labels on peptides of the same kind, we were able to account for this broadening and specifically measure the distance between a spin-label on an hr peptide and on a bt peptide. To do this, the spectrum of hr.m/bt.wt was chosen as the noninteracting spectrum because hr.m/bt.wt is a 3-spin system in which all broadening of the EPR spectrum is due to interactions of peptides of the same kind. The choice of noninteracting spectrum was assumed to be the dominant source of error. Experimental error was estimated by calculating the distances using the other complexes containing three spin-labels (hr.n/bt.wt, hr.c/bt.wt, hr.wt/bt.Ca, and hr.wt/bt.Cb) as the noninteracting spectrum (Table 1). When the spectrum of AK6 (Figure 4) was used as the noninteracting spectrum for the calculations, slightly shorter distances were obtained. The distances measured for hr.m/bt.Ca and hr.n/bt.Cb were both 3% shorter. All other distances were at most 9% shorter. Thus, regardless of the specifics of the calculations, the rankings are the same.

RESULTS

A Helical Trimer of Heterodimers. Peptides including most of the heptad repeat region (hr) and peptides from closer to the transmembrane domain (bt) of gp41 were synthesized (Figure 1). The peptide hr.wt is the same as dp107 (Wild *et al.*, 1992), and its membrane binding characteristics were studied previously (Rabenstein & Shin, 1995a). These peptides are also part of the protease-resistant core of gp41; hr.wt is the N-terminal 38 residues of N-51, and bt is the C-terminal 34 residues of C-43 (Lu *et al.*, 1995). The name "hr.wt" refers to the heptad repeat (hr) region with wild-type (wt) sequence. The names "hr.n, hr.m, and hr.c" refer to the heptad repeat region spin-labeled near the N-terminus, in the middle, and near the C-terminus of the peptide,

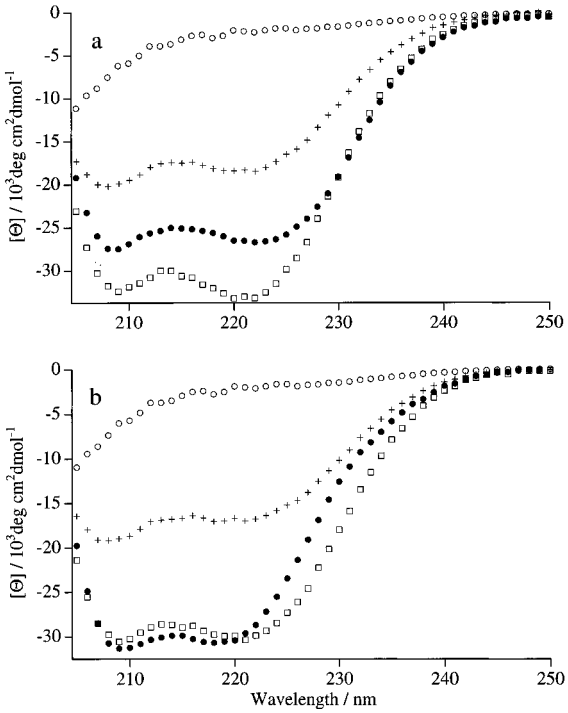

FIGURE 2: (a) CD spectra of hr.wt (boxes), bt.wt (circles), and hr.wt/bt.wt mixed 1:1 (filled circles) and the theoretical spectrum for noninteracting hr.wt and bt.wt (crosses). (b) CD spectra of spin-labeled hr.c (boxes), bt.wt (circles), and spin-labeled hr.c/bt.wt mixed 1:1 (filled circles) and the theoretical spectrum for non-interacting spin-labeled hr.c and bt.wt (crosses). The spectra of hr.wt/bt.wt and hr.c/bt.wt are different from the spectra expected if the peptides did not interact. The mean residue ellipticity of hr.wt/bt.wt is less than that of hr.wt. However, because the hr.wt/bt.wt complex has 72 residues while hr.wt has 38, the total ellipticity, and thus the number of helical residues, is greater in hr.wt/bt.wt.

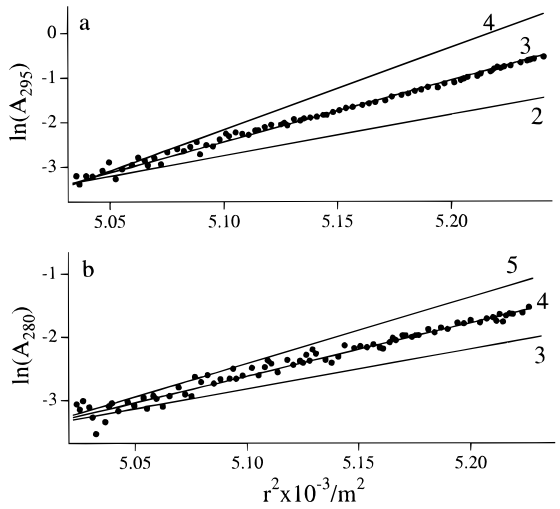

FIGURE 3: Determination of the molecular weight of the hr.wt/bt.wt complex and hr.wt by sedimentation equilibrium ultracentrifugation. (a) Typical data along with calculated data for a dimer (2), trimer (3), and tetramer (4) of hr.wt/bt.wt dimers. The initial concentrations used were 4, 8, 16, and 32 μM of each peptide. b) Data for hr.wt, along with calculated data for a trimer (3), tetramer (4), and pentamer (5) of hr.wt. The initial concentration was 20 μM.

respectively. The name "bt.wt" refers to a peptide of wild-type sequence thought to buttress against the heptad repeat. The name "bt.Ca" refers to bt.wt with a Cys substitution near the N-terminus, and "bt.Cb" refers to bt.wt with a Cys substitution near the middle of the peptide. The secondary structure of hr/bt mixtures was studied by circular dichroism (CD) spectroscopy, the oligomerization by analytical ultracentrifugation, and the tertiary structure by electron paramagnetic resonance (EPR) spectroscopy.

hr.wt and bt.wt interact and form a helical complex. CD spectroscopy revealed that the peptide hr.wt has a mean molar ellipticity at 222 nm ([θ_{222}]) of −33 000 (Figure 2) at 10 μM and 22 °C, indicating that this peptide is approximately 100% helical (Chen *et al.*, 1974). bt.wt has a CD spectrum characteristic of a random coil, with [θ_{222}] of −5000 (Figure 2). When hr.wt and bt.wt peptides are mixed 1:1, the CD spectrum of the mixture is not the sum of the individual spectra (Figure 2). If the two peptides did not interact, [θ_{222}] would be −19 800 (average of the values for bt.wt and hr.wt correcting for the difference in number of residues), but it was measured to be −26 400, consistent with approximately 80% α-helix (Chen *et al.*, 1974). This indicates that α-helical structure is induced by mixing.

hr.wt and bt.wt form a trimer of heterodimers. The oligomerization of the hr.wt/bt.wt complex was studied by equilibrium sedimentation ultracentrifugation. When hr.wt and bt.wt were mixed 1:1, the complex was found to form a trimer of hr.wt/bt.wt dimers (Figure 3). The peptide hr.wt alone was previously found to be tetrameric (Rabenstein & Shin, 1995a). In case the oligomerization is buffer-dependent, we repeated the equilibrium sedimentation experiments of hr.wt under the same buffer conditions used for hr.wt/bt.wt, and again found that hr.wt is tetrameric (Figure 3). Shugars *et al.* (1996) found that 49 residues of the leucine zipper region of gp41 can cause the normally monomeric maltose binding protein to tetramerize, consistent with our results.

Model of the hr/bt Complex Determined by EPR. In order to study the tertiary structure of the hr/bt complex, peptides with Cys substitutions were synthesized (Figure 1) and spin-labeled with 3-(2-iodoacetamido)-PROXYL, a thiol-specific spin-label. Spin-labels within 25 Å have measurable dipolar interactions. In a frozen sample, dipolar interactions cause a broadening in the EPR spectrum dependent on the distance between the spin-labels. The closer the spin-labels, the broader the spectrum. The distance can be quantitatively measured (Rabenstein & Shin, 1995b). Qualitative measurements can be made at ambient temperatures (Fiori *et al.*, 1994).

The EPR spectra of all combinations of hr and bt are shown in Figure 4. Qualitatively, the spectrum of the mixture of hr.n/bt.Cb is the broadest, indicating the spin-labels on hr.n/bt.Cb are closer than any other pair. The spectrum of hr.m/bt.Ca is the next broadest, followed by hr.m/bt.Cb, hr.n/bt.Ca, and hr.c/bt.Ca. The spectrum of hr.c/bt.Cb is not broader than those of hr.c/bt.wt and hr.wt/bt.Cb (Figure 4), indicating that the spin-labels on hr.c do not interact with the spin-labels on bt.Ca. Because no interaction is observed, the separation of the spin-labels is greater than ∼25 Å. The distances between spin-labels were also quantitatively determined (Table 1) using the method described by Rabenstein and Shin (1995b).

The distances measured are graphically presented in Figure 5. The Cys mutations hr.n and hr.m are separated by 10 residues, corresponding to a 15 Å translation in an α helix. The Cys substitutions in hr.n and hr.m are at the "c" and "f" positions, respectively, so they are on the same face but are

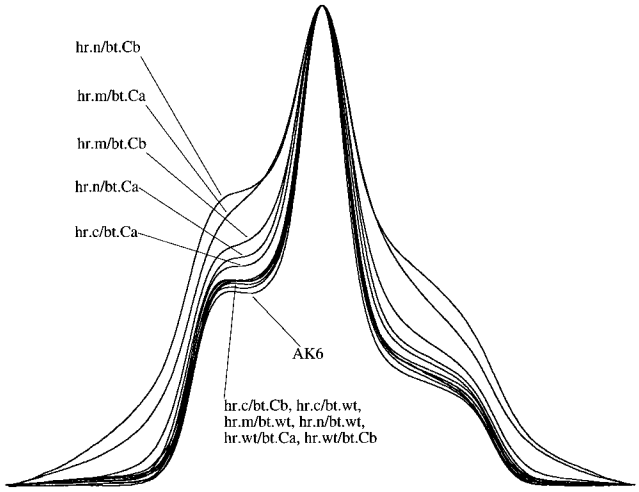

FIGURE 4: Integrated EPR spectra of spin-labeled hr/bt complexes at −140 °C. The sweep width was 150 G. When the spin-label on the hr peptide interacts with the spin-label on the bt peptide, the spectrum broadens. The broader the spectrum, the closer the spin-labels. AK6 is the spin-labeled monomeric peptide Ac-AAAAK-CAAAKAAAAKAAAAKA-NH_2 (Marqusee *et al.*, 1989).

Table 1: Measured Distances (Å) between Spin-Labels[a]

	hr peptide		
bt peptide	hr.n	hr.m	hr.c
bt.Ca	15.6 ± 0.9	10.3 ± 0.3	17.4 ± 1.5
bt.Cb	9.5 ± 0.3	13.4 ± 0.5	>25

[a] Distances and errors were calculated as described under Materials and Methods.

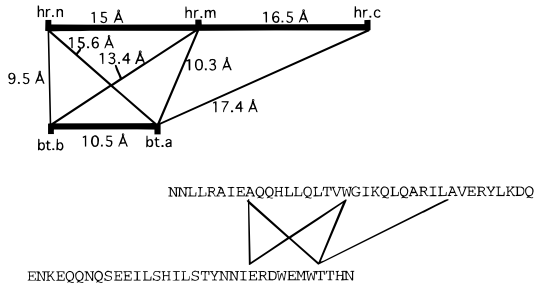

FIGURE 5: Measured distances in relation to peptide secondary structure. The intrapeptide distances were calculated assuming the peptides are α-helical with 1.5 Å between residues.

not completely aligned. The Cys in hr.c is also at the "c" position, and 11 residues from the Cys mutation in hr.m, which is a 16.5 Å translation. The Cys mutations bt.Ca and bt.Cb are separated by 7 residues and are both at the "g" position. This corresponds to a 10.5 Å separation.

In the model, the distance between the spin-labels on hr.n and bt.Ca is less than the distance between the spin-labels on hr.m and bt.Cb. However, the distances measured were 15.6 and 13.4 Å, respectively. This discrepancy may be due to several factors. First, there is some uncertainty in the measured distances (Table 1). Second, the hr.m spin-label is at the "f" position, while the hr.n spin-label is at the "c" position of the heptad repeat. Although they are on the same face, they are not completely aligned. Finally, the spin-label side chain is flexible and exists in several conformations. Due to steric factors, certain conformations may be preferred. Thus, the distance between two spin-labels might not exactly reflect the distance between the residues to which they are attached.

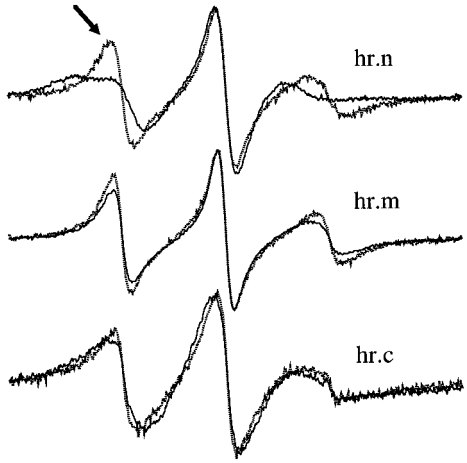

FIGURE 6: EPR spectra of hr.n, hr.m, and hr.c with (solid lines) and without (dashed lines) bt.wt at 22 °C. The spectra of hr.m and hr.c show little change, indicating the motion of the spin-labels was not greatly affected by the addition of bt.wt. The spectrum of hr.n, however, broadens, indicating a decrease in mobility of the spin-label. These spectra are consistent with bt.wt packing against the spin-labels on hr.n, but not hr.c and hr.m.

Because the exact orientations and conformations of the spin-labels are not known, only translation along the helix is considered. For this study, exact distances are not as important as the observation that the hr.m spin-label is closer to the bt.Ca spin-label than to the bt.Cb spin-label. Furthermore, the hr.n spin-label is closer to the bt.Cb spin-label than to the bt.Ca spin-label. These qualitative observations alone indicate that the helices are antiparallel and not completely overlapping, as depicted in Figure 5.

Tertiary Contacts Support the Model. The spectrum of a spin-label on a peptide at room temperature is a function of the rotational correlation time (τ_c) of the spin-label, anisotropy in the motion, and interactions with proximate spin-labels. The EPR spectrum is especially sensitive to τ_c's in the range 0.1−100 ns (Millhauser *et al.*, 1995). Furthermore, spin-labels are sensitive to tertiary interactions (Altenbach *et al.*, 1990).

At 22 °C, when hr.m or hr.c was mixed with bt.wt, the EPR spectrum did not significantly change (Figure 6). However, when hr.n was mixed with bt.wt, there was a large change in the EPR spectrum, indicative of a slower τ_c. This suggests that bt.wt contacts hr at the position of the Cys mutation in hr.n, but not at the positions of the Cys mutations in hr.m or hr.c. This is consistent with bt.wt not completely overlapping hr.wt, as suggested by the distances measured (Figure 5). The spectrum of the hr.n/bt.wt complex was not dependent on the concentration. Furthermore, a similar change was not observed in either hr.m or hr.c. This indicates that the change in the spectrum of hr.n upon addition of bt.wt is not due to nonspecific aggregation.

A Helical Complex with Greater Overlap. CD spectra of another peptide derived from the heptad repeat of gp41 support the alignment. hr2.wt is a 38-residue peptide with its sequence derived from the heptad repeat region of gp41, shifted 13 residues toward the N-terminus with respect to the peptide hr.wt. According to the model presented above, the N-terminus of hr2.wt is aligned with the C-terminus of bt.wt. The sequence of hr2.wt is Ac-QARQLLSGIV-QQQNNLLRAIEAQQHLLQLTVWGIKQLQ-NH_2. The CD spectrum of hr2.wt showed no signs of helicity, but when mixed with bt.wt, the CD spectrum became characteristic

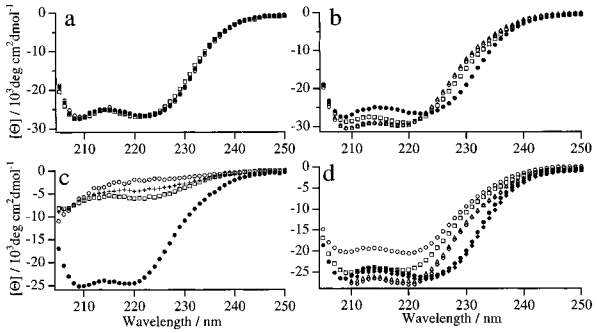

FIGURE 7: (a) CD spectra of hr.wt mixed 1:1 with bt.wt (filled circles), bt.Ca (squares), and bt.Cb (open circles). (b) CD spectra of hr.wt/bt.wt (filled circles), hr.n/bt.wt (open circles), hr.m/bt.wt (squares), and hr.c/bt.wt (triangles). (c) CD spectra of hr2.wt (boxes), bt.wt (circles), hr.wt/bt.wt mixed 1:1 (filled circles), and the theoretical spectrum for noninteracting hr.wt and bt.wt (crosses). (d) CD spectra of hr.wt/bt.wt at 5 μM (filled circles) and 100 μM (filled diamonds), and of hr2.wt/bt.wt at 5 μM (open circles), 10 μM (open squares), 50 μM (open triangles), and 100 μM (open diamonds).

of an α-helix (Figure 7). At concentrations at which the spectra were concentration-independent (>50 μM), the amplitude of the hr2.wt/bt.wt spectrum from 210 to 222 nm was greater than that of hr.wt/bt.wt (Figure 7). This suggests that the hr2.wt/bt.wt complex has more helicity than hr.wt/bt.wt, consistent with the alignment determined by EPR.

CD Spectra of Spin-Labeled Peptides. In order to determine perturbations caused by the Cys mutations and spin-labels, the CD spectra of all peptide mixtures were measured. As with hr.wt/bt.wt, the CD spectrum of hr.c/bt.wt shows interaction between the two peptides consistent with bt.wt becoming somewhat helical (Figure 2). hr.m/bt.wt and hr.n/bt.wt showed analogous results (Figure 7), although the spectral shapes are not identical to that of hr.wt/bt.wt. The minimum at 222 nm was shifted to 220 nm, and there was approximately a 5% increase in the mean residue ellipticity. The CD spectrum, however, is still characteristic of an α-helix, indicating the structure is not significantly altered by the spin-labels.

The CD spectra of hr.wt mixed with each of the bt peptides (both wild-type and spin-labeled) are virtually identical (Figure 7). Similarly, spectra of hr.n, hr.m, and hr.c mixed with bt.Ca or bt.Cb are nearly identical to spectra of each of these peptides mixed with bt.wt (data not shown). Furthermore, the mean residue ellipticity was independent of the peptide concentration in the range 5−100 μM (Figure 7).

DISCUSSION

The hr.wt/bt.wt complex has high helical content, as determined by CD spectroscopy. At 10 μM, $[\theta_{222}] = -26\,400$ for hr/bt, corresponding to approximately 80% helicity. Increasing the concentration to 100 μM did not significantly change the CD spectrum (Figure 7), indicating the fraction not oligomerized was insignificant. Thus, 80% helicity probably means that 80% of the residues are helical all the time, and approximately 58 out of 72 residues (hr.wt and bt.wt) are helical. Assuming all 38 residues of hr.wt are helical, bt.wt is 59% helical. This corresponds to 20 helical and 14 non-helical residues in bt.wt. From the model, 15 residues extend beyond the edge of the hr peptide; the 14 nonhelical residues are most likely at the C-terminus of the bt peptide.

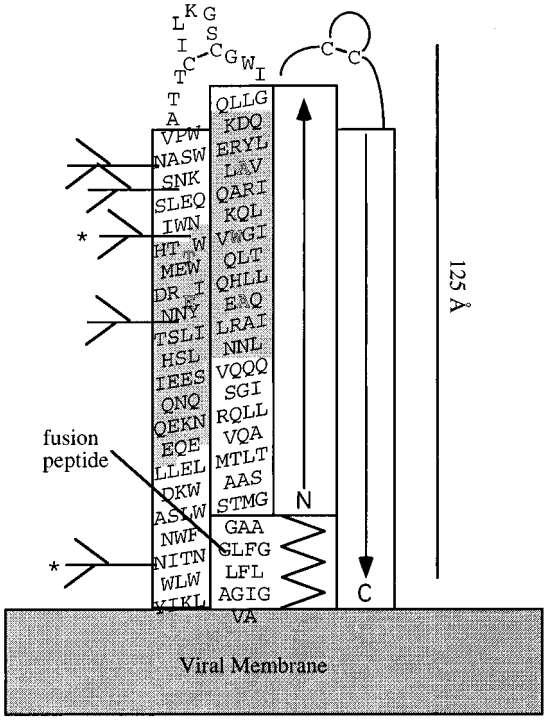

FIGURE 8: Proposed model of gp41 showing two subunits of a trimer. Helices are shown as rectangles, with the sequence written from left to right. The shaded residues are hr.wt and bt.wt. The residues that were spin-labeled are shown in outline font. Glycosylation and a disulfide bond are also indicated. Nonconsensus glycosylation is indicated with an asterisk. The distance 125 Å is based on 1.5 Å/ residue for an α-helix, and is consistent with electron micrographs of the ectodomain, excluding the fusion peptide (AASMT...FNITN, 146 residues), showing rods of length 125 ± 13 Å.

Using the alignment shown in Figure 5, 11 residues of C43 extend the N-terminus of N51 in the N51/C43 complex. If these 11 residues were nonhelical, 88% of the C43/N51 complex would be helical. In fact, C43/N51 was measured to be approximately 89% helical (Lu *et al.*, 1995), consistent with our alignment.

A second peptide, hr2.wt, has a more dramatic change in the CD spectrum when mixed with bt.wt. According to our model, bt.wt aligns with hr2.wt with no residues extending beyond the N-terminus of hr2.wt. Contrasting hr.wt, hr2.wt is not helical without bt.wt. However, the hr2.wt/bt.wt complex is slightly more helical than the hr.wt/bt.wt complex at 100 μM. For hr2.wt/bt.wt, $[\theta_{222}] = -27\,300$, corresponding to approximately 83% helicity, or 12 nonhelical residues. Because hr2.wt is not helical without bt.wt, it is not reasonable to assume that hr2.wt is 100% helical within the hr2.wt/bt.wt complex. The 12 nonhelical residues may be from either peptide, and are probably due to "fraying" on the ends. For example, the C-terminal 3 residues of the GCN4 leucine zipper are not helical (O'Shea *et al.*, 1991). Thus, a total of 12 residues at the ends of hr2.wt and bt.wt may be unstructured, leading to 83% helicity. The increase from 0 to 83% helicity in unmixed and mixed hr2.wt/bt.wt complex suggests that they are well-aligned, and that the alignment based on the EPR data is correct.

gp41 Model. Based on these results and previous studies, we propose a model for the structure of the ectodomain of gp41 (Figure 8). We propose that following the fusion peptide, there is a trimeric coiled-coil approximately 67 residues long, followed by a disulfide bond containing a

β-turn. This is followed by a second α-helical region, approximately 73 residues long, which continues into the transmembrane region.

The overall shape and secondary structure are consistent with those determined by electron microscopy by Weissenhorn *et al.* (1996). The ectodomain, excluding the fusion peptide (AASMT...FNITN, 146 residues), was observed to be a rod 125 ± 13 Å long. This is consistent with the length of the helices in our model. Furthermore, the 146-residue protein expressed was 80% α-helical by CD. In our model, this region is 91% helical. However, according to our model, 14 residues (10%) of the protein fragment studied by Weissenhorn *et al.* (1996) extend over the end of a central coiled-coil. These residues may be uncoiled in the protein fragment, although they are helical in intact gp41. Alternatively, perhaps less of gp41 is helical than we propose, as there is not much information about the structure of the top of the protein (see below).

Although we found that approximately 12 residues of bt.wt are nonhelical in the hr.wt/bt.wt complex, this region is probably helical in full-length gp41. As discussed previously, the 12 nonhelical residues of bt.wt are probably a consequence of their not being in contact with a core coiled-coil.

Several studies indicate that the hr region is a coiled-coil, with helices "buttressing" against it. A peptide from this region has been found to form a tetrameric coiled-coil on its own (Wild *et al.*, 1992; Rabenstein & Shin, 1995a). Peptides from the buttressing helix are unstructured on their own, but become α-helical when mixed with peptides from the buttressing helix (Lu *et al.*, 1995). Previously, no direct information was available concerning the alignment of the helices. In this study, we determined the relative positions of the inner coiled-coil and the buttressing helices.

We propose that gp41 is trimeric because this study and two others of ectodomain fragments containing both the bt and hr regions found a trimeric oligomerization state (Lu *et al.*, 1995; Weissenhorn *et al.*, 1996).

The immunodominant region (residues 589−604) contains two essential and conserved Cys residues. An oxidized peptide from this region was found to form a type-one reverse turn (Oldstone *et al.*, 1991). In our model, this region is the furthest from the viral membrane, consistent with its being the immunodominant region.

The precise C-terminal end of the central coiled-coil is unclear. We propose that the coiled-coil is terminated at Gly-594. The beginning of the buttressing helix is also unclear. We propose that it begins 16 residues before the protease-resistant core of the bacterially expressed gp41 ectodomain (Lu *et al.*, 1995). The first 17 residues of the proposed buttressing helix contain 3 glycosylation sites. Glycosylation at these sites may be important for the structure of this region. When expressed in *Escherichia coli*, proteins are not glycosylated, and thus may fold incorrectly, leading to susceptibility to proteolysis. We propose that Pro-609 is the second residue of the buttressing helix because Pro occurs frequently at the second position in an α-helix (Rose & Wolfenden, 1993). This extends the buttressing helix nearly to the top of the central coiled-coil.

This model is consistent with the glycosylation sites on gp41. Glycosylation occurs on Asn residues with the signal sequence Asn/X/(Thr or Ser) where X is any residue except Pro. The conserved glycosylation sites all lie on the outer buttressing helix. This helix also contains an imperfect heptad repeat of hydrophobic residues which likely pack against the inner coiled-coil. The glycosylation sites occur at "e," "c," and "b" positions, and not at the core "a" and "d" positions. There is an additional, nonconserved glycosylation site at residue 624 of the HXB2R isolate, which is position "d". The consensus sequence for HIV-1, however, has a glycosylation site at the next residue, which is at the "e" position (Myers *et al.*, 1993). Thus, glycosylation of the consensus sequence is consistent with this model, although the HXB2R isolate is an exception.

The N-terminal fusion peptide is potentially near the viral membrane, approximately 100 Å from the top of the ectodomain. It is not clear what the secondary structure of the fusion peptide is. However, even if it extends away from the viral membrane, the immunodominant region is still the furthest part of the peptide from the viral membrane. Because the fusion peptide is likely the first part of gp41 to interact with the cellular membrane, this structure may represent the prefusogenic form of gp41. Previously, it has been suggested that the ectodomain of gp41 without gp120 should form the fusogenic form of gp41 (Lu *et al.*, 1995; Weissenhorn *et al.*, 1996). This assumes, however, that gp41 undergoes a transition from a prefusogenic form to a fusogenic form upon CD4 binding, with the fusogenic form being the thermodynamic minimum, and the prefusogenic form a kinetically-trapped metastable state. Furthermore, it assumes that folded sections of gp41 are at the same thermodynamic minimum as when the rest of gp41 (including the hydrophobic fusion peptide), gp120, and the viral membrane are present. Finally, this two-state model is based on studies of influenza hemagglutinin (HA) (Baker & Agard, 1994). Because of many differences between gp41 and HA, HA may not be a good model for gp41.

Comparison with HA2. Like gp160 of HIV, HA is the envelope glycoprotein of the influenza virus responsible for receptor binding and viral−cellular membrane fusion. HA is composed of HA1, the receptor binding subunit, and HA2, the TM domain. HA2 is the most-studied viral membrane fusion protein. It has similarities in sequence and function to gp41, which lead to an early model of gp41 based on HA2 (Gallaher *et al.*, 1989). For example, both are viral TM fusion proteins formed by proteolytic cleavage of a larger polypeptide, and both contain heptad repeats. However, because of significant differences between these proteins, comparisons should be made cautiously.

Although the tertiary structure of gp41 is not known, there is evidence that it is different from HA2. At neutral pH, the N-terminal fusion peptide of HA2 is followed by (among other structures) a helix which buttresses against a subsequent 50-residue coiled-coil (Wilson *et al.*, 1981). In contrast, the N-terminal fusion peptide of gp41 is followed by a coiled-coil, which is followed by a buttressing helix. Thus, gp41 is "inside-out" relative to HA2; the N-terminal region forms a coiled-coil, with a region only 18 residues from the transmembrane domain buttressing against it.

The disulfide bond locations of gp160 and HA are also different. In contrast to gp120 and gp41, which are noncovalently linked, HA1 and HA2 are linked by a disulfide bond. Furthermore, gp41 has two conserved cysteines separated by five residues which are thought to form a disulfide bond. HA2 has a similar disulfide bond. However,

in gp41, it is 80 residues from the transmembrane domain, while in HA2 it is only 37 residues from the transmembrane domain.

Additionally, gp41 and HA2 are functionally different. Although they both mediate viral−cellular membrane fusion, HA2 is triggered by a decrease in pH when the virus is endocytosed. However, neither endocytosis nor a pH change is required for HIV viral entry.

Here we present data on the tertiary structure of gp41. In conjunction with previous studies, we present a model for the structure of gp41 which is significantly different from HA2. Further experiments will clarify similarities and differences between HA2 and gp41.

ACKNOWLEDGMENT

We thank Thorgeir E. Thorgeirsson for the use of his EPR data analysis programs, D. S. King of the Howard Hughes Medical Institute for mass spectrometry, Ralph Peteranderl for help with ultracentrifugation, and Anne Fu and Kyubyul Hwang for help with peptide synthesis.

REFERENCES

Altenbach, C., Marti, T., Khorana, H. G., & Hubbell, W. L. (1990) *Science 248*, 1088−1092.

Baker, D., & Agard, D. A. (1994) *Biochemistry 33*, 7505−7509.

Chen, Y. H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.

Delwart, E. L., Mosialos, G., & Gilmore, T. (1990) *AIDS Res. Hum. Retroviruses 6*, 703−706.

Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.

Fiori, W. R., Lundberg, K. M., & Millhauser, G. L. (1994) *Nat. Struct. Biol. 1*, 374−377.

Gallaher, W. R., Ball, J. M., Garry, R. F., Griffin, M. C., & Montelaro, R. C. (1989) *AIDS Res. Hum. Retroviruses 5*, 431−440.

Lu, M., Blacklow, S. C., & Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 1075−1082.

Marqusee, S., Robbins, V. H., & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286−5290.

McMeekin, T. L., Groves, M. L., & Hipp, N. J. (1949) *J. Am. Chem. Soc. 71*, 3298−3300.

Millhauser, G., Fiori, W., & Miick, S. (1995) in *Methods in Enzymology* (Sauer, K., Ed.) pp 589−610, Academic Press, New York.

Myers, G., Koorber, B., Wain-Hobson, S., Smith, R. F., & Pavlankis, G. N. (1993) *Human Retroviruses and AIDS*, Los Alamos National Laboratory, Los Alamos, NM.

Oldstone, M. B., Tishon, A., Lewicki, H., Dyson, H. J., Feher, V. A., Assa, M. N., & Wright, P. E. (1991) *J. Virol. 65*, 1727−1734.

O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539−544.

Parry, D. A. (1982) *Biosci. Rep. 2*, 1017−1024.

Rabenstein, M., & Shin, Y. K. (1995a) *Biochemistry 34*, 13390−13397.

Rabenstein, M. D., & Shin, Y. K. (1995b) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8239−8243.

Rose, G. D., & Wolfenden, R. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 381−415.

Shugars, D. C., Wild, C. T., Greenwell, T. K., & Matthews, T. J. (1996) *J. Virol. 70*, 2982−2991.

Weast, R. C. (1975) *Handbook of Chemistry and Physics*, CRC Press, Cleveland.

Weissenhorn, W., Wharton, S. A., Calder, L. J., Earl, P. L., Moss, B., Aliprandis, E., Skehel, J. J., & Wiley, D. C. (1996) *EMBO J. 15*, 1507−1514.

Wild, C., Oas, T., McDanal, C., Bolognesi, D., & Matthews, T. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10537−10541.

Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature 289*, 366−373.

BI961743T