Biophysical Characterization of the Structure of the Amino-terminal Region of gp41 of HIV-1

IMPLICATIONS ON VIRAL FUSION MECHANISM*

(Received for publication, September 23, 1998)

Ding-Kwo Chang‡, Shu-Fang Cheng, and Vishwa Deo Trivedi
From the Institute of Chemistry, Academia Sinica, Taipei, 11529 Taiwan, Republic of China

A peptide of 51 amino acids corresponding to the NH₂-terminal region (5–55) of the glycoprotein gp41 of human immunodeficiency virus type 1 was synthesized to study its conformation and assembly. Nuclear magnetic resonance experiments indicated the sequence NH₂-terminal to the leucine zipper-like domain of gp41 was induced into helix in the micellar solution, in agreement with circular dichroism data. Light scattering experiment showed that the peptide molecules self-assembled in water into trimeric structure on average. That the peptide molecules oligomerize in aqueous solution was supported by gel filtration and diffusion coefficient experiments. Molecular dynamics simulation based on the NMR data revealed a flexible region adjacent to the hydrophobic NH₂ terminus of gp41. The biological significance of the present findings on the conformational flexibility and the propensity of oligomerization of the peptide may be envisioned by a proposed model for the interaction of gp41 with membranes during fusion process.

Fusion between enveloped viruses and host cells is an essential step for viral infection (1). The fusion domain within the transmembrane (TM) proteins has been shown to be responsible for virus-cell fusion (2). For a majority of viruses, the fusion peptide, which penetrates into the target membrane, has been found at the NH₂ terminus of the TM subunit (3). A conformational change in the fusion domain of the virus, triggered by low pH (4) or binding to the receptors (5) leading to exposure of the fusion peptide, has been proposed to be an important intermediate step for virus-mediated fusions (for review, see Ref. 6). As suggested by many experiments, the fusion proteins oligomerize in the fusogenic state (7), and the fusion pore is believed to be composed of several oligomeric units of fusion proteins (8–10). Thus, accumulation of the virus glycoproteins has also been found at the cell-cell contact region (11). Recently an early fusion active conformation of HIV-1 TM gp41 has been detected using an inhibitory synthetic peptide (12).

Recent structural studies by x-ray crystallography on the hemagglutinin of influenza virus (7) and the TM glycoprotein gp41 of human immunodeficiency virus revealed that the core of the fusion domain consists of an inner triple-stranded coiled-coil buttressed by three helices formed by the COOH-terminal region of the ectodomain of the envelope glycoproteins (13–15). An anti-parallel core complex of the fusion domain of fusion-mediating virus has also been deduced from a recent solution NMR study on a truncated protein derived from the ectodomain of simian immunodeficiency virus (16). The common features of these helix bundles are that the central coiled-coil trimer is formed by the leucine zipper-like heptad repeat sequence and that the height of the cylinder of the complexes is at least 60 Å. Thus, it is of interest to examine how the fusion complex containing several long rigid cylinders initiates approach of the merging membranes.

The function of the NH₂-terminal fusion peptide of gp41 has been characterized, and the ensuing more polar region has been shown to associate with the surface subunit, gp120, of the external domain of HIV-1 env glycoprotein. These two regions of gp41 have also been implicated in the viral fusion by mutational studies (17). Despite elucidation of the core of the viral fusion domain complex by the x-ray diffraction and NMR studies, the structural role in the fusion event of the fusion peptide and the intervening sequence between it and the zipper-like region has not been addressed.

To better define the role of a given domain in the functions of a protein, a peptide corresponding to that segment can be studied (18–20). In particular, the NH₂-terminal hydrophobic peptides have been found to cause cell lysis, to induce electric current, and to promote lipid mixing in liposomes (21, 22). A peptide corresponding to amino acids 5 through 16 of gp41 has been found to induce current under external applied voltage (21). The interaction with SDS micelle, on the atomic scale, of the 23-amino acid peptide corresponding to the NH₂-terminal region of gp41 has been studied recently (23). It was found that Ala-Gly dipeptide, located 15 and 16 residues from the NH₂ terminus of gp41, is at the micellar-aqueous boundary and the region COOH-terminal to it is flexible.

In view of the highly conserved nature of the NH₂-terminal region of gp41 implying its important structural role in the fusion (24), we report here the structure and conformations of a peptide encompassing amino acids 5 through 55 of gp41, under varying solution conditions. Using NMR, circular dichroism spectroscopies, light scattering, and gel filtration chromatography we attempt to further unravel the virus-mediated cell fusion in light of the reported oligomeric core structure of gp41 (13, 14, 16). The NMR methods are particularly suited for our purposes because they afford atomic resolution, the observation of conformational change in aqueous and membrane-mimic environments, and information regarding the flexibility of the specific regions in the peptides or proteins under study. In combination with the crystalline structural data on the core complex of gp41 external domain, the results show the physical
properties for the NH2-terminal region of the TM glycoprotein gp41 and suggest the conformational transition during the fusion event and the hemifusion intermediate. Similar changes in conformation and assembly state may also occur in the fusion process of other enveloped viruses such as influenza virus.

EXPERIMENTAL PROCEDURES

Materials—The peptide (formyl-GALFLGFLGAAAGSTMGRSMTLTVQARQLSGIVQNNLRAIEAEQHLLL-OH, gp41-N51) corresponding to amino acids 5–55 of gp41 and the peptide (NH2-YTFSLISHLIESQNOQEQKNEQELLEDKWSLLWNWFOH, gp41-C36) corresponding to amino acids 127–162 from strain LAV– (LAV, lymph adenopathy-associated virus) were synthesized in an automated mode by a solid phase synthesizer from Applied Biosystems (Foster City, CA) model 431A using Fmoc (N-9-fluorenylmethoxycarbonyl) chemistry. The peptides were cleaved from the resin by trifluoroacetic acids and purified by reverse phase high-performance liquid chromatography using a C8 column. The primary sequence of the peptide was ascertained by electrospray mass spectrometry as well as amino acid analysis.

SDS (sodium dodecyl sulfate) was acquired from Boehringer Mannheim (Mannheim, Germany) and d25-SDS from Cambridge Isotope (Andover, MA). 1,2-Ditetredecylsn-glycer-3-phosphocholine (PC) and 1,2-dihexadecanoyl-sn-glycer-3-phospho-t-serine (PS) were acquired from CalBiopeptide (San Diego, CA). Lysozyme (molecular weight d kDa) was purchased from Sigma. All reagents were used in the experiments without further purification. Solutions containing vesicles were prepared by solubilizing the lipids in chloroform:methanol (4:1, v/v) sonicated for at least 0.5 h before measurements. The NOE interactions for the peptide in SDS micellar solution (molecular mass: 81, 43, 13.7, and 6.5 kDa, respectively).

Gel Filtration Chromatography—Measurements were performed in a Hitachi L-6000 liquid chromatography using a Superdex-75 HR 10/30 column (Pharmacia). The peptide and protein samples of 0.4–0.8 mg/ml concentrations were eluted with solution containing 22% acetonitrile and 0.1% trifluoroacetic acid with a flow rate of 0.5 ml/min. Data were recorded at a wavelength of 214 nm. Transferrin, ovalbumin, ribonuclease A, and aprotinin were used as molecular mass markers (molecular mass: 81, 43, 13.7, and 6.5 kDa, respectively).

NMR Data Indicate Enhanced Helical Structure for gp41-N51 in the Presence of SDS Micelles—Sequential assignment of 1H NMR chemical shift of gp41-N51 in the SDS micellar dispersion was accomplished from the fingerprint (Fig. 1A) and NH/NH regions (Fig. 1B) to resolve overlapping cross-peaks. The 1H chemical shift assignment of the peptide in aqueous solution devoid of the detergent dispersion was made with the same strategy. The NOE interactions in these two solutions are summarized in Fig. 1, C and D (chemical shifts of 1H resonances for the peptide in SDS micellar solution and in water are tabulated in Table I). Resonance peaks of the NH2-terminal region of the TM glycoprotein gp41 of HIV-1 was accomplished by fitting CD data with Hennessey-Johnson algorithm (25), using 33 proteins of known secondary structure as the basis set.

NMR Data Indicate Enhanced Helical Structure for gp41-N51 in the Presence of SDS Micelles—Sequential assignment of 1H NMR chemical shift of gp41-N51 in the SDS micellar dispersion was accomplished from the fingerprint (Fig. 1A) and NH/NH regions (Fig. 1B) to resolve overlapping cross-peaks. The 1H chemical shift assignment of the peptide in aqueous solution devoid of the detergent dispersion was made with the same strategy. The NOE interactions in these two solutions are summarized in Fig. 1, C and D (chemical shifts of 1H resonances for the peptide in SDS micellar solution and in water are tabulated in Table I). Resonance peaks of the pH of Gly3, Gly6, Gly12, and Gly16 in SDS micellar solution are split, suggesting a certain degree of structural rigidity in the NH2-terminal region of the peptide. In contrast, in water pH of glycines residues, with the exception of Gly6, give rise to degenerate chemical shifts, implying more flexible conformation for the peptide in aqueous solution. This is likely due to the fact

RESULTS

Characterization of Amino-terminal Domain of gp41 of HIV-1

A

The fingerprint region (A) and NH/NH region (B) of NOE spectrum of gp41-N51 in SDS micellar solution and in water (D). The leucine zipper region starts with Ile41. Potential NOE peaks due to spectral overlap are designated by the dotted line. The NOE peaks are more sparse in the region Ala12–Leu22, d(NOJ, i) denotes the NOE cross-peak arising from NH of residue i and NH of residue j, d(NOJ, i) denotes the NOE cross-peaks of the side chain protons of residues i with any protons from residue j.

Fig. 1. The fingerprint region (A) and NH/NH region (B) of NOE spectrum of gp41-N51 in SDS micellar solution and in water (D). The leucine zipper region starts with Ile41. Potential NOE peaks due to spectral overlap are designated by the dotted line. The NOE peaks are more sparse in the region Ala12–Leu22, d(NOJ, i) denotes the NOE cross-peak arising from NH of residue i and NH of residue j, d(NOJ, i) denotes the NOE cross-peaks of the side chain protons of residues i with any protons from residue j.
| Residue | NH        | αH        | βH         | Others |
|---------|-----------|-----------|------------|--------|
| Gly ^1  | 8.66      | 3.99 3.93 | 3.63       |        |
| Ala ^2  | 8.71      | 4.33      | 1.40       |        |
| Leu ^3  | 8.47      | 4.32 1.57 | 1.77 1.75  |        |
| Phe ^4  | 8.55      | 4.66 3.08 | 2.6H       | 7.22   |
| Leu ^5  | 8.41      | 4.31 1.54 | γH 1.51    | δCH3 0.92 0.86 |
| Gly ^6  | 7.41      | 3.82 3.72 | 3.91       |        |
| Phe ^7  | 8.25      | 4.66 3.14 | 2.6H 7.22  | 7.21   |
| Leu ^8  | 8.46      | 4.32 1.57 | γH 1.49    | δCH3 0.92 0.86 |
| Gly ^9  | 7.96      | 3.87      | 3.91       |        |
| Ala ^10 | 8.34      | 4.32 1.42 | 1.32       |        |
| Ala ^11 | 8.57      | 4.31 1.47 | 1.32       |        |
| Gly ^12 | 8.50      | 3.97      | 3.90 3.90  |        |
| Ser ^13 | 8.31      | 4.55 3.97 | 3.90 3.90  |        |
| Thr ^14 | 8.40      | 4.41 4.31 | γCH3 1.23  |        |
| Met ^15 | 8.52      | 4.49 2.12 | γCH2 2.65  | 2.57   |
| Gly ^16 | 8.56      | 3.94      | 3.94 3.94  |        |
| Ala ^17 | 8.32      | 4.31 1.42 | 1.32       |        |
| Arg ^18 | 8.50      | 4.35 1.89 | γCH2 1.68  |        |
| Ser ^19 | 8.45      | 4.43 3.90 | 3.86 3.96  |        |
| Met ^20 | 8.61      | 4.59 2.16 | γCH2 2.67  | 2.58   |
| Thr ^21 | 8.31      | 4.31 4.26 | γCH3 1.23  |        |
| Leu ^22 | 8.44      | 4.47 1.70 | γH 1.67    |        |
| Thr ^23 | 8.27      | 4.34 4.34 | γCH3 1.27  |        |
| Val ^24 | 8.49      | 4.39 4.09 | γCH3 1.33  |        |
| Gln ^25 | 8.69      | 4.23 2.14 | γCH2 2.43  |        |
| Ala ^26 | 8.22      | 4.30 1.47 | 1.51       |        |
| Characterization of Amino-terminal Domain of gp41 of HIV-1 |
|----------------------------------------------------------|
| **TABLE I** continued |  |
| **NH** | **αH** | **βH** | **γCH₂** | **δCH₂** | **δNH₂** | **Others** |
| Arg27 | 8.44 | 4.18 | 1.881.73 | 1.68 | 1.64 |  |
| | 8.23 | 3.92 | 2.031.96 | 1.80 | 1.68 |  |
| | 3.23 | 3.25 |  |
| Gln28 | 8.31 | 4.32 | 2.052.00 | 2.43 | 2.38 |  |
| | 7.94 | 4.11 | 2.252.20 | 2.52 | 2.46 |  |
| | 7.64 | 7.04 |  |
| Leu29 | 8.29 | 4.25 | 1.74 | 1.63 | 1.72 |  |
| | 8.04 | 4.23 | 1.891.76 | 0.89 | 0.94 |  |
| | 0.98 | 0.95 |  |
| Leu30 | 8.22 | 4.37 | 1.75 | 1.63 | 1.63 |  |
| | 8.22 | 4.12 | 1.86 | 0.90 | 0.91 |  |
| |  |
| Ser31 | 8.31 | 4.41 | 3.963.93 |  |
| Gly32 | 8.45 | 4.00 |  |
| Ile33 | 8.03 | 4.12 | 1.90 | 1.53 | 1.19 |  |
| | 7.78 | 4.01 | 2.05 | 1.78 | 1.29 |  |
| |  |
| Val34 | 8.34 | 3.99 | 2.07 | 1.00 | 0.95 |  |
| Gln35 | 8.56 | 4.26 | 2.04 | 2.41 | 2.48 |  |
| | 7.92 | 4.25 | 2.272.16 | 2.56 | 2.48 |  |
| | 7.42 | 6.79 |  |
| Gln36 | 8.52 | 4.27 | 2.162.12 | 2.44 | 2.31 |  |
| | 7.87 | 4.37 | 2.17 | 2.50 | 2.31 |  |
| | 7.68 | 6.76 |  |
| Gln37 | 8.61 | 4.26 | 2.222.14 | 2.43 | 2.31 |  |
| | 8.08 | 4.38 | 2.04 | 2.48 | 2.31 |  |
| | 7.67 | 7.01 |  |
| Asn38 | 8.65 | 4.67 | 2.86 | 7.76 | 7.07 |  |
| Asn39 | 8.17 | 4.61 | 2.922.86 | 7.61 | 6.78 |  |
| | 8.60 | 4.62 | 2.87 | 7.76 | 7.07 |  |
| | 8.44 | 4.59 | 2.89 | 7.65 | 6.91 |  |
| Leu40 | 8.22 | 4.28 | 1.75 | 1.67 | 1.58 |  |
| | 8.10 | 4.17 | 1.801.73 | 0.94 | 0.90 |  |
| | 0.99 | 0.93 |  |
| Leu41 | 8.09 | 4.27 | 1.75 | 1.65 | 1.71 |  |
| | 8.05 | 4.05 | 1.881.75 | 0.97 | 0.88 |  |
| | 0.98 | 0.93 |  |
| Arg42 | 8.20 | 4.24 | 1.90 | 1.74 | 1.66 |  |
| | 7.93 | 4.02 | 1.921.83 | 1.69 | 1.66 |  |
| | 3.24 | 3.24 |  |
| |  |
| Ala43 | 8.31 | 4.28 | 1.46 | 1.62 | 1.22 |  |
| | 7.86 | 4.22 | 1.58 | 1.86 | 1.18 |  |
| | 1.86 | 1.18 |  |
| Ile44 | 8.20 | 4.01 | 1.92 | 1.62 | 1.22 |  |
| | 8.09 | 3.80 | 2.03 | 1.86 | 1.18 |  |
| | 0.94 | 0.94 |  |
| Glu45 | 8.45 | 4.20 | 2.09 | 2.45 | 2.37 |  |
| Ala46 | 8.28 | 4.04 | 2.18 | 2.45 | 2.37 |  |
| Gln47 | 8.18 | 4.24 | 2.11 | 2.45 | 2.37 |  |
| | 7.73 | 4.30 | 2.282.11 | 2.53 | 2.42 |  |
| | 7.66 | 6.99 |  |
| Gln48 | 8.44 | 4.26 | 2.11 | 2.45 | 2.37 |  |
| | 7.92 | 4.24 | 2.12 | 2.45 | 2.37 |  |
that Gly⁶ is in the FLG motif shown to form a type I β-turn in the fusion peptide of gp41 in aqueous solution (27).

To identify the secondary structure, the chemical shifts are compared with the random coil values as compiled by Wishart and Sykes (28), and the result is presented in Fig. 2. In water (Fig. 2B), the helical structure can be detected in the Gln 25–Leu 30 stretch and is more prominent in the region COOH-terminal to Gln 35. This result is in agreement with the NOE results shown in Fig. 1D where δ(i,i+3) interactions, characteristic of helical form, are observed in these regions. In the presence of SDS micelles (Fig. 2A), on the other hand, the regions adopting helical structure encompass Leu²⁻Gly¹⁵, Val²⁴⁻Val³⁴, and Asn³⁸⁻Leu⁵⁰. The sequence following Ile⁴⁴ has the zipper 4-3 repeat pattern and thus is expected to have high propensity of forming helical structure. The NH₂-terminal fusion peptide of gp41 has been shown to insert into membraneous structure primarily as a helix (23). In addition to the helix formation of the NH₂-terminal fusion peptide in SDS solution, enhancement of the helical form in the micellar medium is also observed for the region Gln²⁵⁻Ser³¹ and the region following Gln³⁵. This is evident from the larger chemical shift deviation of δH and βH from the random values in SDS solution than those obtained in water as shown in Fig. 2. The present data for gp41-N51 thus corroborate with the idea that helix is induced in the micellar environment.

CD Data Also Reveal Enhancement in α-Helix for gp41-N51 in Vesicular and Micellar Solutions—The CD results displayed in Fig. 3 are consistent with the NMR data in that the helical content for the peptide is significantly increased upon association with SDS micelle. The helical structure is also enhanced in both PS and PC vesicular solutions. From the CD data, the secondary structural element analyzed by the algorithm of Hennessey and Johnson (25) for gp41-N51 is summarized in Table II. The result clearly indicates an increase in the helix.

|     | NH  | αH  | βH  | δNH₂ | 2H  | 4H  |
|-----|-----|-----|-----|------|-----|-----|
| His⁴⁹ | 8.53 | 4.74 | 3.333.21 | 7.65 | 7.99 |
|     | 8.32 | 4.75 | 3.352.6 | 8.03 | 8.66 |
| Leu⁵⁰ | 8.44 | 4.40 | 1.65 | γH  | 7.38 |
|     | 8.10 | 4.34 | 1.861.74 | 1.63 | 0.94 |
|     | 0.97 | 0.90 |
| Leu⁵¹ | 8.03 | 4.23 | 1.65 | γH  | 1.61 |
|     | 7.93 | 4.10 | 1.67 | δCH₃ | 0.93 |
|     | 0.95 | 0.91 |
population in the membrane-mimic environment, especially in dispersions that carry negatively charged headgroups.

**Oligomerization of gp41-N51 in Water Is Indicated by Light Scattering Experiments**—The results of light scattering measurements on gp41-N51 and lysozyme are presented in Fig. 4. The intercept for gp41-N51 is 5.71 × 10⁻⁸, which yields a particle mass of 17.5 kDa. The value is approximately three times the monomeric mass of 5.4 kDa for gp41-N51. The molecular weight determined by light scattering experiments is a weight-averaged value, hence the data indicate that the peptide molecules are predominantly trimeric or they are in an equilibrium between dimers and tetramers. The error estimated from the difference between the experimental and reported values of lysozyme is 9%.

**Gel Filtration Experiment and Diffusion Coefficient Measurement Support Oligomerization of gp41-N51 Molecules in Aqueous Solution**—The gel filtration data on gp41-N51 displayed in Fig. 5A may arise from dimerization of the peptide. Oligomerization of the NH₂-terminal peptide of gp41 is also demonstrated by diffusion coefficient measurements from NMR pulsed field gradient experiments shown in Fig. 5B. The D₀ of the peptide is slightly larger than that of lysozyme with a molecular mass of 14.4 kDa. Thus, as in the light scattering experiments, diffusion coefficient measurements indicate that the peptide molecules exist in the aqueous solution as trimers, or there is equilibrium between dimers and tetramers for the peptide. It is of interest to observe that a particle mass of about 15 kDa is obtained for the peptide in the concentration on the order of 1 mg·ml⁻¹ for light scattering and NMR experiments, while smaller size is detected at lower concentration employed in gel filtration experiment. The oligomerization of the NH₂-terminal peptide of shorter length encompassing 33 amino acids has also been reported by Kliger et al. (29).

**Three-dimensional Structures of gp41-N51 Derived from NOE Data Suggest a Flexible Segment Adjacent to the Hydrophobic NH₂-terminal Region**—The three-dimensional structure of gp41-N51 in the presence of SDS micelles was calculated by employing the constraints derived from NOE data and hydrogen bonds suggested by the deuteron-hydrogen exchange experiments. Fig. 6A displays a typical computed structure and the superposition of 10 structures excluding residues between Ala¹⁰ and Val²⁴. As expected, the helical form can be observed in the NH₂-terminal region up to Ala¹¹, and the segments Leu²²-Val³⁴ and Asn³⁵-Leu⁵⁰. The sequence between Ala¹¹ and Val²⁴ exhibits substantial flexibility as indicated by Fig. 6B, which shows larger root mean square deviation values of backbone atoms for these residues than residues in the remainder of the peptide as analyzed from the 10 conformers used in Fig. 6A. This flexible region overlaps with the gp120-associating site implicated in the mutagenesis studies (1).

The root mean square deviation values of atomic fluctuation for Asn³⁵ is substantially larger than those of the adjacent residues. In addition, a 3₁₀ helix is found for the Gln³⁵-Gln³⁷ stretch (Φ(35) = -71 ± 9°, Ψ(35) = -30 ± 7°, Φ(36) = -82 ± 11°, Ψ(36) = -37 ± 14°, and Φ(37) = -79 ± 22°, Ψ(37) = -35 ± 13°), which is readily transformed into α-helix. The structures are thus convergent except for the segment span-

![Image](71x177 to 275x394)

**FIG. 3.** Far-UV-CD spectrum of gp41-N51 in water (——) and in SDS micellar (–––), PS-PC vesicular (molar ratio of peptide: PS:PC = 1:75:25, ······· = 1:25:75, ——) solutions. It appears that the helicity of the peptide is higher in the presence of negatively charged micelle or vesicle.

![Image](327x203 to 535x438)

**FIG. 4.** Light scattering data on gp41-N51 (×) and lysozyme (○, molecular mass: 14.4 kDa) in water. The intercepts of KcR₀ versus sin²(θ/2), representing the reciprocal of the average molecular mass, are 5.71 × 10⁻⁸ and 6.37 × 10⁻⁸, respectively, for gp41-N51 and lysozyme at room temperature. The data thus yield molecular masses of 17.5 ± 1.4 and 15.7 ± 1.0 kDa, respectively, for gp41-N51 and lysozyme, indicating that the NH₂-terminal portion of gp41 exists in aqueous solution approximately as trimers.

**TABLE II**

| Solution | Helix | Antiparallel β-sheet | Parallel β-sheet | Turn | Others |
|----------|-------|----------------------|------------------|------|--------|
| H₂O      | 10 ± 2| 19 ± 2               | 5 ± 1            | 30 ± 1| 36 ± 1 |
| SDS      | 51 ± 1| 1 ± 1                | 3 ± 0            | 23 ± 1| 22 ± 0 |
| PS:PC = 75:25 | 60 ± 2| 9 ± 1                | 8 ± 1            | 4 ± 1 | 19 ± 1 |
| PS:PC = 25:75 | 26 ± 3| 45 ± 8               | 2 ± 3            | 2 ± 2 | 25 ± 9 |

**Secondary structure analysis from CD data of gp41-N51 in aqueous, SDS, and lipidic solutions (in %)**

Data were obtained from CD spectra shown in Fig. 3.
ning Ala\textsuperscript{11}–Leu\textsuperscript{22}. Aside from both ends of gp41-N51 that are subject to the fraying effect, the larger fluctuation is seen for the Ala\textsuperscript{11}–Leu\textsuperscript{22} region. The result concurs with the idea that the Ala\textsuperscript{11}–Gly\textsuperscript{12} dipeptide is localized at the micellar-aqueous boundary, as noted previously (23), and determines the depth of penetration for the fusion peptide into the micelle. As will be discussed in the next section, the flexibility of this segment may have implication in the fusogenic activity of the NH\textsubscript{2}-terminal portion of gp41.

**DISCUSSION**

**Conformational Change of the Hydrophobic Fusion Peptide in Membrane-mimic Environment and the Flexibility of the Adjacent Polar Region of gp41 Suggest an Essential Role for These Two Regions in the Virus-mediated Fusion Process**—Our CD data in water, the micellar and vesicular solutions for both the 23-amino acid fusion peptide sequence of gp41 in a previous report (23), and a longer gp41-N51 peptide indicate that, in the membranous environment, the helix form is induced at the expense of \( \beta \)-form or random coil for the NH\textsubscript{2}-terminal hydrophobic stretch of gp41. As shown previously (23), and determines the depth of penetration for the fusion peptide into the micelle. As will be discussed in the next section, the flexibility of this segment may have implication in the fusogenic activity of the NH\textsubscript{2}-terminal portion of gp41.

**适应性ability of gp41-N51 on the amino acid level exhibited in the presence and absence of membrane-mimic environment is provided by NMR data. However, conformational of the more polar sequence following the fusion peptide, Ala\textsuperscript{12}–Leu\textsuperscript{22}, was found more flexible from structural analysis of NMR data (Fig. 6). This conclusion is also consistent with scarcity of NOE cross-peaks involving protons in the residues of Ala\textsuperscript{12}–Leu\textsuperscript{22} segment, likely due to the smaller motional correlation times and/or more disordered structure, both of which are compatible with the idea of a flexible region. Because the segment was found to be essential for association with gp120 (2, 30), the flexibility of this region in the membranous medium suggests that it plays a role in the conformational change during the native-to-fusogenic transformation of gp41, which is likely triggered by the dissociation of gp120 from gp41 upon binding to CD4 and the secondary receptor (for example, CCR5 or CXCR4).

As indicated in Figs. 2, A and B, the FLG motif (Phe\textsuperscript{4}–Gly\textsuperscript{6}) undergoes a \( \beta \) to \( \alpha \) transition as the peptide interacts with the SDS micelle. Moreover, the motif has been shown to serve as an
FIG. 7. Proposed model for the interaction of gp41 with target and viral membranes during fusion process. Top panel, the core oligomeric coiled-coils associating with the COOH-terminal helices of gp41 ectodomain. The NH2-terminal fusion peptide is primarily in β- and random forms and the sequence between the fusion peptide and zipper-like domain, which may have a loose structure, is not completely dissociated from gp120 and/or the co-receptor on the target membrane. This is due to the highly hydrophobic nature of the fusion peptide, which does not favor its complete exposure to the polar solvent. Initial contact of gp41 with the target membrane is probably through the FLG motif. Middle, complete dissociation of gp41 from other components of fusion complex exposes the fusion peptide, which subsequently inserts into, and destabilizes, both the target and viral membranes primarily as a helix. Helical structure is also enhanced for the sequence NH2-terminal to the zipper-like region. According to the determined depth of penetration of fusion peptide, only the outer leaflet of the membrane bilayer of target cell has direct contact with the viral fusion peptide. Driven probably by the oligomeric coiled-coil zipper-like domain, the NH2-terminal region of the gp41 molecules self-associate. Propensity of the NH2-terminal region of gp41 to oligomerize, in turn, induces bulged surfaces on the two approximate membranes when the fusion peptides insert into these two membranes, thus facilitating the membrane merger. Alternatively, the interaction of the flexible region of gp41-N51 with the base region of the gp41 ectodomain can promote membrane fusion. Note that some hydrophobic NH2-terminal sequences may insert into the viral membrane. The straddling of the NH2 termini from a core complex of gp41 over the target and viral membranes and the tendency to oligomerize by the NH2-terminal regions of gp41 provide the energy required to dehydrate the membranes to be merged. As a consequence, the two opposing membranes are brought into contact and dehydration occurs as the two membranes begin to attach. From the figure it can be visualized that the importance of the pliability of the segment Ala11-Leu22 to allow the core coiled-coil rod-like structure with length
initiation site for helix formation (27). Together with insertion into the micelle of the NH₂-terminal Gly¹–Ala¹⁰ region, as suggested by the spin label attenuation data (23), these results suggest that the FLG motif may make initial contact with, and penetrate into, the membrane as part of the conformational change of gp41 following binding of the HIV envelope glycoprotein to its receptors in the fusion event (31, 32).

The Ala¹¹-Gly¹² region where the structures begin to diverge as shown in Fig. 6B coincides with the site found at the micellar-aqueous interface from spin label attenuation experiment for the fusion peptide in SDS solution. These results are reasonable in that the segment of the peptide in the micellar interior is more restricted than the region lying on the external surface of the micelle. Since the hydrocarbon chains of the cellular membrane are generally longer than that of SDS micelle, our data indicate that the inner leaflet of the target membrane bilayer is not likely to be reached by the viral fusion peptide during fusion. Therefore, it is probably that the direct effect of viral fusion peptide is to cause the hemifusion, as first suggested by White and co-workers on the fusion mediated by hemagglutinin of the influenza virus (33).

The increase in helicity for the segment Gln²⁵–Ser⁴¹ and the stretch Ser¹⁰–Val²⁴ (Fig. 2) in the presence of SDS micelles strongly suggests that these changes are part of the structural transition in gp41 as it transforms from fusion-inactive to fusion-active state. As already mentioned, the Gln³⁵–Asn³⁸ stretch is likely to be helical in the fusogenic state. The enhancement of helical form for gp41-N51 in the membranous medium suggests that helix is the dominant form during membrane fusion.

Oligomerization of gp41-N51 in Water Detected by Light Scattering, Gel Filtration, and Diffusion Coefficient Measurements Indicates the Propensity of Self-assembly of the Peptide and Suggest a Functional Role of Merging Viral and Target Cell Membranes for the NH₂-terminal Portion of gp41—The oligomerization state of the envelope glycoprotein of HIV-1 has been examined by various techniques, which yielded differing results of dimeric, trimeric, and tetrameric structures (13, 14, 34). It is, however, agreed that the membrane-anchoring subunit gp41 is primarily responsible for the oligomerization. In the present work, the peptide corresponding to the NH₂-terminal region of gp41 was found to be oligomeric in water. Although the mass values deduced from gel filtration and diffusion coefficient data depend somewhat on the molecular shape, the experiments point to a non-monomeric state for gp41-N51. However, more definitive result on the oligomerization is provided by light scattering data which give the absolute molecular mass from the intercept of a typical Zimm plot (26, 35). Dimers are apparently present in the more dilute solution for the gel filtration experiment, whereas trimers or mixtures of dimers and tetramers can account for data obtained in the more concentrated solution for NMR diffusion coefficient measurements. The latter conclusion is in agreement with the result obtained from the light scattering experiment in the same concentration range. Hence it is probable that different association states found in various studies arise from different conditions under which experiments are performed. Different oligomerization and assembly states might also be present at various stages of fusion event. Since gp41-N51 peptide molecules oligomerize in aqueous environment, it is likely that the NH₂-terminal portion of gp41 penetrates into the target membrane in oligomeric form during some stage of fusion process.

Analogous examples for the zipper-like domain to induce oligomerization have been reported for the protein A/gp41(538–593) chimeric protein (36) and for the maltose-binding protein/gp41(558–595) chimeric protein (34). The zipper-like domain encompassing amino acids 555–583 of gp41 is responsible for oligomerization of the chimeric proteins, since both protein A and maltose-binding protein are monomeric (34, 36). The stability of oligomerization is apparently critical in merging the phospholipids of the two apposing membranes. This is supported by the observation that a mutation in gp41 fusion peptide dominantly interfered with fusion and infectivity (37), which can be rationalized by participation of the defective mutant protein in the oligomer, rendering the fusion complex ineffective.

It has been shown by a number of investigators that the peptide corresponding to the hydrophobic NH₂-terminal fusion peptide sequence is sufficient to cause leakage and phospholipid mixing of the vesicular bilayers (22, 38). The vesicle and micelle are more prone to fuse than planar cellular membrane because of their smaller radius of curvature (39). Thus, to effect the fusion of planar bilayers of the viral and cellular membranes, it is possible that the core oligomeric complex formed by the zipper-like coiled-coil and COOH-terminal helices of gp41 ectodomain acts as the joint of a molecular nipper to hold the oligomeric structure, while the fusion peptide segments that penetrate into the membranes and the flexible polar border sequences would bring the two opposing membranes into contact by acting as the arms of a nipper. More explicitly, since the extended rod-like cylinder is at least 60 Å in length, the helical length would be extended to more than 90 Å if the intervening sequence between zipper-like domain and the hydrophobic fusion peptide also assumes rigid rod conformation.
making approach of the two apposing membranes impossible. The flexibility of the Ala^{11}–Leu^{22} stretch, illustrated in Fig. 6, A and B, and the propensity of oligomerization of the NH_{2}-portion of gp41 render possible direct contact of phospholipids of the opposing membranes (Fig. 7). This is done by bending the Ala^{11}–Leu^{22} segment and the base of the gp41 ectodomain to make the core oligomeric complex parallel to the membrane surfaces. Flexibility of the sequence near the base of gp41 ectodomain was suggested by the crystallographic study of the NH_{2}- and COOH-terminal helices of gp41.

Energetically, a driving force to merge the two approaching membranes is provided by the tendency for the NH_{2}-terminal portion of gp41, represented by gp41-N51, to self-assemble and/or by the interaction between the NH_{2}-terminal region and the COOH-terminal region of the gp41 external domain. The latter hypothesis is based on the fluorescence data on gp41-C36 incubated with gp41-N51 (see Fig. 8). The model illustrated in Fig. 7 emphasizes the essential role in bringing the two fusing membranes into approximation played by the segment NH_{2}-terminal to the zipper-like domain of gp41. Because the core antiparallel helix bundle of gp41 functions as the joint or hinge of a nick, it cannot exert force on the target or viral membranes. The latter function is then furnished by the arm of the nicker, the polar intervening sequence (Ala^{11}–Leu^{22}) of gp41, making approach of the two apposing membranes impossible. This is done by bending the flexible arms (Ala^{11}–Leu^{22} of gp41-N51), due to the tendency to oligomerize for the NH_{2}-terminal portion of gp41 extending from the core structure and/or the base region of the gp41 ectodomain can promote membrane coalescence, as illustrated in Fig. 7. It is stressed that the fusion peptide, because of its hydrophobicity, is not likely to be completely exposed to solvent during the entire fusion process. The fusion peptide may be ushered by other component(s) of the fusion reaction, such as gp120 or the receptor or the co-receptor, to the target membrane.

Membrane fusion is an energy costing event, since it involves destabilization of the bilayers and dehydrogenation of the interface regions of merging membranes (40). This notion is also consistent with the low effectiveness of productive entry by the virus into the target cell (41). The insertion of the hydrophobic fusion peptide makes initial perturbation of the target membrane. Propensity of the region encompassing gp41-N51 for oligomerization and the interaction between the flexible, more polar NH_{2}-terminal region outside the membranes and the segment preceding the membrane-anchoring sequence of gp41 probably helps to overcome the hurdle of membrane dehydrogenation during the membrane merging.

In summary, study of the NH_{2}-terminal region of gp41 complements the structural investigation of core complex formed by the oligomers of the zipper-like domain and the COOH-terminal helix. The findings of flexibility of the Ala^{11}–Val^{24} segment and the tendency of forming oligomers by gp41-N51 and its interaction with the region near the base of the ectodomain of gp41 along with increased helicity in the membranous medium for the peptide suggests an structural role of the fusogenic activity of the NH_{2}-terminal portion of gp41. Our data also imply a stalk and a hemifusion intermediate for the fusion reaction mediated by HIV. Similar mechanism may be used by other enveloped viruses in mediating cell fusion.

Acknowledgments—We are grateful to Dr. Shui-Tein Chen for synthesis of the peptides and Dr. Ming-Fai Tam for characterization of the peptides by mass spectrometry and to Dr. Wei-Jyun Chien for acquisition of some NMR data.

REFERENCES

1. Hernandez, L. D., Hoffman, L. R., Wolfsberg, T. G., and White, J. M. (1996) Annu. Rev. Cell. Dev. Biol. 12, 627–661
2. Cao, J., Bergerman, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. (1993) J. Virol. 67, 2747–2755
3. White, J. M. (1990) Annu. Rev. Physiol. 52, 675–697
4. White, J. M., Matlin, K., and Helenius, A. (1981) J. Cell Biol. 89, 674–679
5. Sattentau, Q. J., and Moore, J. P. (1990) J. Exp. Biol. 174, 407–415
6. Chan, D. C., and Kim, P. S. (1998) Cell 93, 681–684
7. Builough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) Nature 371, 57–63
8. Spruce, A., Iwata, A., and Almers, W. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 3623–3627
9. Ellens, H., Bentz, J., Mason, D., Zhang, F., and White, J. M. (1990) Biochemistry 29, 8097–8099
10. White, J. M. (1992) Science 258, 917–924
11. Aroeti, B., and Henis, Y. I. (1991) J. Biol. Chem. 266, 15845–15849
12. Furuta, R. A., Wild, C. T., Weng, Y., and Weiss, C. D. (1998) Nat. Struct. Biol. 5, 276–279
13. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) Cell 89, 263–273
14. Weisshaarborn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) Nature 387, 426–430
15. Tan, K., Liu, J.-H., Wang, J. H., Shen, S., and Lu, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 12503–12508
16. Chaffey, M., Cai, M., Kaufmann, S. J., Stahl, S. J., Wingfield, P. T.,Cogli, D. V., Gronenborn, A. M., and Clore, G. M. (1998) EMBO J. 17, 4572–4584
17. Freed, E. O., Myers, D. J., and Rissel, R. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4650–4654
18. O’Shea, E. K., Rutikowski, R., and Kim, P. S. (1989) Science 243, 538–542
19. Wild, C., Dubay, J. W., Greenwell, T., Baird, T., Jr., Oas, T. G., McDanal, C., Hunter, E., and Matthews, T. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 12676–12680
20. Pristker, M., Jones, P., Blumenthal, R., and Shai, Y. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 7287–7292

Characterization of Amino-terminal Domain of gp41 of HIV-1

5309