Published in final edited form as:

Angew Chem Int Ed Engl. 2013 April 2; 52(14): 3911–3915. doi:10.1002/anie.201207266.

Internal Dynamics of the Homotrimeric HIV-1 Viral Coat Protein gp41 on Multiple Time Scales**

Dr Nils-Alexander Lakomek

Laboratory of Chemical Physics, NIDDK National Institutes of Health Building 5, Room 126, 9000 Rockville Pike Bethesda, MD 20892-0520, U.S.A.

Joshua D. Kaufman and Dr Stephen J. Stahl

Protein Expression Laboratory, NIAMS National Institutes of Health Building: 6B, Room: 1B130

Dr John M. Louis and Dr Alexander Grishaev

Laboratory of Chemical Physics, NIDDK National Institutes of Health Building 5, Room 126, 9000 Rockville Pike Bethesda, MD 20892-0520, U.S.A.

Dr Paul T. Wingfield

Protein Expression Laboratory, NIAMS National Institutes of Health Building: 6B, Room: 1B130

Dr Ad Bax*

Laboratory of Chemical Physics, NIDDK National Institutes of Health Building 5, Room 126, 9000 Rockville Pike Bethesda, MD 20892-0520, U.S.A.

Keywords

NMR; membrane proteins; HIV; gp41; protein dynamics

Fusion of viral and cellular membranes is elicited by the HIV-1 envelope glycoprotein gp120/gp41. The precursor gp160, encoded by the *Env* gene, is cleaved post-translationally into two chains, gp120 and gp41, which remain non-covalently associated as a homotrimer of heterodimers and form a spike on the viral surface.^[1] Upon CD4 recognition, the gp120 subunit dissociates from gp41, which remains anchored through its C-terminal transmembrane helix (TM) in the viral membrane. After gp120 dissociation, the N-terminal fusion peptide of gp41 is exposed and can insert into the host cell membrane.^[2] Because of its high sequence conservation and its accessibility to the humoral immune system,^[3] gp41 represents an attractive drug target for antiviral therapy as well as a key player in vaccine research.^[4, 5]

In the current model, membrane fusion is driven by a conformational change of the gp41 ectodomain from an extended pre-fusion (or so-called pre-hairpin) intermediate during virus to host cell docking to a trimer of hairpins, forming an anti-parallel six-helical bundle (6HB) arrangement, thereby pulling viral and host cell membranes to close juxtaposition and

**We thank Yang Shen for expert advice and Jinfa Ying and James Baber for experimental assistance. This work was funded by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH) and the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH. We acknowledge use of the shared scattering beamline resource allocated under the PUP-77 agreement between the National Cancer Institute and the Argonne National Laboratory and thank Dr. Soenke Seifert (Argonne National Laboratory) for his support of the SAXS experiments. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-ENG-38.

*Fax: (+1301-496-2848) bax@nih.gov.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

initiating hemifusion.^[6, 7] This model is mainly based on the X-ray crystal studies of the soluble gp41 ectodomain which revealed the N- and C-terminal heptad repeat regions (NHR and CHR) to pack together as a tight anti-parallel 6HB,^[8, 9] representing the late-fusion or post-fusion conformation of the gp41 ectodomain.^[6, 10] The fusion mechanism itself is believed to be similar to the much better analyzed hemagglutinin system from influenza,^[11, 12] but no structural information on any pre-fusion or early-fusion intermediate of gp41 is yet available. While the structures of the N-terminal fusion peptide (FP) and the membrane proximal external region have been studied extensively by solution and solid state NMR as well as by EPR spectroscopy,^[13–18] structural information on full length gp41 has remained elusive.

Here, we present a solution NMR study on the structure and dynamics of the homotrimeric gp41 complex, encompassing residues 1–194 (512–705 in the numbering of the Env precursor, see Figure 1) reconstituted in dodecyl phosphatidylcholine (DPC) micelles. Li and Tamm showed very similar CD spectra for the gp41 fusion peptide embedded in DPC micelles and in POPC:POPG (4:1) lipid bilayers, indicating the same helicity. The fusion peptide embedded in liposomes composed of the same lipid ratio also induced lipid mixing.^[14] Sun et al. observed very similar structures of the MPER peptide in DPC micelles and in DHPC/DMPC bicelles, with their EPR spectra also showing similar results for the MPER peptide in DPC and in virus membrane-like liposomes.^[16] Therefore, DPC micelles appear to be a suitable mimic for a membrane environment.

Sedimentation equilibrium centrifugation and size exclusion chromatography with in-line multi-angle light scattering and refractive index measurements both show a monomer-trimer equilibrium with a K_D in the low μ -molar range that will be fully shifted towards the homotrimer at NMR concentrations. The total mass of the trimer-detergent complex is 181 kDa (see Supporting Information).

Only ca. 55% of the expected amide correlations are visible in the TROSY-HSQC spectrum (Figure 2) and in the corresponding HNCO spectrum. Observed resonances were assigned to the FP, NHR and immuno dominant loop (IL) regions, whereas CHR, MPER and TM domains remain completely invisible to solution NMR. This absence of the C-terminal resonances together with the large variations in resonance intensities of the NMR-visible N-terminal part of the protein point to highly non-uniform dynamics. We therefore set out to characterize the backbone motions of the gp41^{1–194} trimer by ¹⁵N relaxation. ¹⁵N R_1 and $R_{1\rho}$ relaxation rates as well as ¹⁵N-¹H NOE and transverse ¹⁵N CSA-dipolar cross-correlated relaxation rates, η_{xy} , were measured at both 600 and 800 MHz, using TROSY-based methods optimized for perdeuterated proteins^[19] (Figure 3a–d). Limited sensitivity did not permit the use of 3D methods,^[20] and only data for amides that could be resolved in the 2D ¹⁵N-¹H TROSY-HSQC spectrum are shown in Figure 3. Conformational changes on the micro- to millisecond time scale give rise to exchange contributions to R_2 . These so-called R_{ex} contributions are best measured on the slowly relaxing ¹⁵N-¹H doublet component, using a Hahn-echo experiment (Figure 3e, for details see Supporting Information, Sections 2.3 and 3.3). Relaxation data show random-coil like behavior for the N-terminal residues V2 and G3 and high mobility for regions A15–L26 and L81–Q110, indicated by smaller $R_{2,0}$ ($R_{1\rho}$) and higher R_1 rates as well as reduced ¹⁵N-¹H NOE values. The most rigid visible residues are located C-terminal in the NHR region (I62 to L76). Relaxation data can be clustered into regions of similar dynamics: I4–L12, Q29–R46, I62–L76, T18–M24, and G86–N105. For a complete overview on error-weighted average values of relaxation data and their standard deviations we refer to Supporting Information Table S2a. Relative to the NHR, the fully α -helical FP (assigned to residue I4 to L12, based on ¹³C $^\alpha$ secondary chemical shifts) shows lower R_2 ($R_{1\rho}$) and higher R_1 rates, indicative of large amplitude rigid body dynamics of the FP relative to the larger ectodomain. The

chemical shifts of the most N-terminal 22 residues of gp41^{1–194} agree very closely with those reported previously¹³ for a short FP construct (residues 1–30) in sodium dodecyl sulfate (SDS) micelles (Figure S9), indicating they adopt very similar structures. Increased mobility and reduced α -helical propensity, indicative of transient helical structure (see Figure S9a), is observed for the region stretching from G13 to L26 (compare Figure 3 and Table S2), pointing to substantial flexibility in this linker region connecting FP and NHR.

The limits of the NHR (29–79) and CHR regions (113–155) have previously been deduced from proteolytic digestion studies^[21] that resulted in proteolysis-resistant soluble gp41 fragments which were used for subsequent crystallization studies.^[9, 22] The proteolytically sensitive residues 80–112, missing in all X-ray studies, were ascribed to a loop region. Indeed, our ¹⁵N relaxation and $\Delta\delta^{13}\text{C}^\alpha$ data indicate the mobile IL region to stretch from L81 to Q110.

To describe the internal dynamics of gp41 and the motion of the FP relative to the NHR region, relaxation data, including the transverse cross-correlated relaxation rate, η_{xy} , were analyzed using the extended model-free approach^[24] as described in detail in the Supporting Information. Fitting of the data was only possible when an internal motion on the nanosecond time scale was invoked.^[25] Internal dynamics are characterized by a generalized order parameter, S^2 , and a correlation time τ , both for the fast, picosecond (S_f^2 , τ_f) and slower, low nanosecond time scale (S_s^2 , τ_s). As summarized in Table 1, all domains of gp41 show a remarkably high degree of internal dynamics on a 2–5 ns time scale, characterized by average order parameters ranging between $S_s^2=0.76$ for the most rigid NMR-visible part of the NHR domain and $S_s^2=0.37$ for the linker region between FP and NHR. The FP itself (I4–L12), with an average order parameter of $S_s^2=0.42$, shows high amplitude motion relative to the NHR region (Q29–R46 and I62–L76). When interpreted in terms of free diffusion of the helical FP in a cone model,^[26] this order parameter corresponds to a cone semi-angle of 42°. The amplitude of the fast motions (on the ps time scale), reflected in $S_f^2=0.77$ is comparable to what is seen in globular proteins, confirming that the FP itself remains a well-ordered helix. By contrast, linker and loop regions T18–M24 and G86–N105 show strongly elevated dynamics on both the ps and ns time scales (Table 1).

As noted above, only about 110 spin systems out of 194 give rise to correlations in the HNCO NMR spectrum, which all belong to either the FP, NHR or IL region, whereas CHR, MPER and TM regions remain unobservable. For the highly hydrophobic TM region, incomplete ¹H^N back exchange after the protein is expressed in D₂O solvent cannot be fully excluded, as this region could remain protected from solvent by a low concentration (below cmc) of detergent used in the purification process. However, for the CHR and MPER regions, this cannot explain the absence of their amide resonances, which therefore must be attributed to strongly increased line widths that result in amplitudes below the observable threshold. Increased line widths are due to elevated ¹H and/or ¹⁵N R_2 relaxation rates, and can result from either slow tumbling in the absence of large internal motions, or from conformational exchange processes on the micro- to millisecond time scale that add an exchange contribution, R_{ex} , to R_2 . While conformational exchange effects seem to be largely absent for the FP, linker and NHR regions, with exception of residues around A22 and N42, a strong exchange contribution is observed for the IL region around A96 (Figure 3e). All exchange effects disappear when using a $T_{1\rho}$ measurement with a spin-lock RF field of 2 kHz (Figure 3b), indicating the observed exchange process must have a time constant much longer than 80 μ s. The observation that the N-proximal region of NHR remains visible while its internal dynamics do not significantly narrow its resonances, together with the existence of a strong conformational exchange effect in the IL region, indicate that the absence of the CHR and MPER regions must be attributed to a conformational exchange process on an intermediate NMR time scale.

The very different relaxation properties of MPER (and presumably TM) compared to FP, which shows the most intense resonances, excludes a significant interaction between FP and TM. Paramagnetic relaxation enhancement (PRE) data recorded for gp41, tagged at the C-terminus (position S192C), show a small but consistent increase of ^1H R_2 rates by $8 \pm 2 \text{ s}^{-1}$ for the N-terminal 25 residues, but no statistically significant change outside this region (Supporting Information Figure S11). The small but consistent increase in ^1H R_2 suggests that the FP and its linker to NHR transiently sample conformations close to the TM region. However, considering how steeply PRE effects scale with distance and how uniform the effects are across the fusion peptide and its linker, contacts between FP and TM never get close or specific, and must have a low population. In all likelihood, the late-fusion stage 6HB, where models predict close proximity between TM and FP, is at most very lowly populated under the conditions of our study.

Small angle X-ray scattering (SAXS) data show a pairwise distribution function with a maximum length vector in the 150–170 Å range (Supporting Information Figure S5). The late-fusion stage 6HB arrangement positions the FP and TM domains at the same end of this bundle and a sole population of this state appears incompatible with the dimensions extracted from the scattering data. The extended pre-fusion three-helical bundle arrangement could reach a length exceeding 200 Å when fully extended, and also is incompatible with the SAXS data. However, with the ~30 residue IL region being dynamically highly disordered, the relative orientation of the NHR and CHR domains will fluctuate. SAXS data are compatible with an ensemble where the CHR helices sample a bundle of orientations with an average inter-helical angle of $\sim 50^\circ$ relative to the NHR (see Supporting Information, Section 1.5).

Thus, ^{15}N relaxation, PRE and SAXS data are compatible with a pre-hairpin intermediate that samples a range of relative CHR vs NHR orientations, possibly in exchange with a low population of the late-fusion 6HB. When transitioning from the pre-hairpin intermediate (Figure 4A) to late-fusion 6HB (Figure 4C), the CHR region is believed to either jackknife and pack against the NHR^[27] or zipper up along the inner NHR coiled-coil.^[28] This would require braking up of the CHR three-helical bundle and at least transient disorder of the CHR and MPER regions during the transition period (Figure 4B). Such a process is anticipated to have a high energy barrier and would dramatically impact chemical shifts of the IL, CHR, MPER and TM regions. The large exchange broadening for CHR and MPER regions observed in our measurements is compatible with such a scenario. Interestingly, the high flexibility of the linker region connecting FP and NHR in the pre-fusion intermediate suggests that in the “harpoon model”,^[9, 22] the main purpose of the FP during the docking stage is to act as an anchor attaching or injecting into the host cell membrane, while stress on the membrane curvature,^[29] at least at this stage, is not transferred from NHR to FP and, if at all, must be generated by the FP itself. The flexibility of the FP-adjacent linker region then allows the re-positioning of the ectodomain when transitioning to the 6HB, driving membrane juxtaposition and subsequent hemifusion stalk formation.

Experimental Section

NMR measurements were carried out on a uniformly $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ -enriched sample of homotrimeric gp41^{1–194} (10 mg/ml, 0.5 mM) in 50 mM sodium acetate buffer, pH 4.0, 25 mM KCl and 330 mM ($\approx 115 \text{ mg/ml}$) DPC, at 313 K. Comparison with a spectrum recorded at pH 7.1 in 50 mM HEPES and 200 mM DPC showed that the HSQC spectrum retains very similar appearance over this pH range (Supporting Information), and all NMR measurements subsequently were conducted at pH 4, where the protein was most stable and spectra were not compromised by amide hydrogen exchange with solvent. Oligomerization of the construct was examined by sedimentation equilibrium ultracentrifugation and size

exclusion chromatography with in-line multi-angle light scattering. NMR backbone assignment was performed as described in the Supporting Information. Backbone dynamics were studied using ^{15}N NMR R_1 , $R_{1\rho}$ relaxation and $\{^1\text{H}\}$ ^{15}N NOE experiments with a TROSY detection scheme^[19] as well as transverse ^{15}N CSA-dipolar cross-correlated relaxation (η_{xy}) measurements. The η_{xy} rates were measured as the difference of the relaxation rate of the slowly and fast relaxing NH doublet component in a TROSY-based experiment.^[30] The presence of conformational exchange effects was studied by a “TROSY-T2” experiment, a simple but informative extension of the regular TROSYHSQC experiment. ^1H R_2 rates were measured to study PRE effects for a S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) spin-labeled S192C mutant of gp41^{1–194} and the reference sample using a TROSY-based detection scheme. Spectra were recorded on a 600 MHz Bruker Avance II, 800 MHz Avance III, and 900 MHz Bruker Avance II system, all equipped with cryo probe head technology. Details on the experimental setup of all experiments can be found in the Supporting Information. Spectra were processed using the NMRPipe/NMRDraw software package^[31]. Secondary chemical shifts have been corrected for ^2H isotope effects using the random coil values of Maltsev et al.^[32] ^{15}N relaxation and η_{xy} rates have been fitted using the extended model-free approach^[24], implemented with a home-written MATLAB script.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- [1]. Zhu P, Liu J, Bess J, Chertova E, Lifson JD, Grise H, Ofek GA, Taylor KA, Roux KH. *Nature*. 2006; 441:847. [PubMed: 16728975]
- [2]. Sattentau Q, Moore J. J. *Exp. Med.* 1991; 174:407. [PubMed: 1713252]
- [3]. Dimitrov DS, Blumenthal R. J. *Virol.* 1994; 68:1956. [PubMed: 8107256]
- [4]. Munch J, Standker L, Adermann K, Schuz A, Schindler M, Chinnadurai R, Pohlmann S, Chaipan C, Biet T, Peters T, Meyer B, Wilhelm D, Lu H, Jing WG, Jiang SB, Forssmann WG, Kirchhoff F. *Cell*. 2007; 129:263. [PubMed: 17448989]
- [5]. Douek DC, Kwong PD, Nabell GJ. *Cell*. 2006; 124:677. [PubMed: 16497577]
- [6]. Harrison SC. *Nat. Struct. Mol. Biol.* 2008; 15:690. [PubMed: 18596815]
- [7]. Harrison SC. *Virus Structure and Assembly*. 2005; Vol. 64:231.
- [8]. Chan DC, Kim PS. *Cell*. 1998; 93:681. [PubMed: 9630213]
- [9]. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. *Nature*. 1997; 387:426. [PubMed: 9163431]
- [10]. Roux KH, Taylor KA. *Curr. Opin. Struct. Biol.* 2007; 17:244. [PubMed: 17395457]
- [11]. Wilson IA, Skehel JJ, Wiley DC. *Nature*. 1981; 289:366. [PubMed: 7464906]
- [12]. Skehel JJ, Wiley DC. *Annu. Rev. Biochem.* 2000; 69:531. [PubMed: 10966468]
- [13]. Jaroniec CP, Kaufman JD, Stahl SJ, Viard M, Blumenthal R, Wingfield PT, Bax A. *Biochemistry*. 2005; 44:16167. [PubMed: 16331977]
- [14]. Li YL, Tamm LK. *Biophys. J.* 2007; 93:876. [PubMed: 17513369]
- [15]. Qiang W, Sun Y, Weliky DP. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:15314. [PubMed: 19706388]
- [16]. Sun ZYJ, Oh KJ, Kim MY, Yu J, Brusica V, Song LK, Qiao ZS, Wang JH, Wagner G, Reinherz EL. *Immunity*. 2008; 28:52. [PubMed: 18191596]
- [17]. Kim M, Sun ZYJ, Rand KD, Shi XM, Song LK, Cheng YX, Fahmy AF, Majumdar S, Ofek G, Yang YP, Kwong PD, Wang JH, Engen JR, Wagner G, Reinherz EL. *Nat. Struct. Mol. Biol.* 2011; 18:1235. [PubMed: 22002224]
- [18]. Lai AL, Moorthy AE, Li YL, Tamm LK. *J. Mol. Biol.* 2012; 418:3. [PubMed: 22343048]

- [19]. Lakomek NA, Ying JF, Bax A. J. Biomol. NMR. 2012; 53:209. [PubMed: 22689066]
- [20]. Caffrey M, Kaufman J, Stahl SJ, Wingfield PT, Gronenborn AM, Clore GM. J. Magn. Reson. 1998; 135:368. [PubMed: 9878465]
- [21]. Lu M, Blacklow SC, Kim PS. Nature Structural Biology. 1995; 2:1075.
- [22]. Chan DC, Fass D, Berger JM, Kim PS. Cell. 1997; 89:263. [PubMed: 9108481]
- [23]. Deleted in proof
- [24]. Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM. J. Am. Chem. Soc. 1990; 112:4989.
- [25]. Korzhnev DM, Orekhov VY, Arseniev AS. J. Magn. Reson. 1997; 127:184. [PubMed: 9281482]
- [26]. Lipari G, Szabo A. J. Am. Chem. Soc. 1982; 104:4546.
- [27]. Blumenthal R, Durell S, Viard M. J. Biol. Chem. 2012; 287:40841. [PubMed: 23043104]
- [28]. Frey G, Rits-Volloch S, Zhang XQ, Schooley RT, Chen B, Harrison SC. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:13938. [PubMed: 16963566]
- [29]. Kozlovsky Y, Kozlov MM. Biophys. J. 2002; 82:882. [PubMed: 11806930]
- [30]. Lee D, Hilty C, Wider G, Wuthrich K. J. Magn. Reson. 2006; 178:72. [PubMed: 16188473]
- [31]. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. J. Biomol. NMR. 1995; 6:277. [PubMed: 8520220]
- [32]. Maltsev AS, Ying JF, Bax A. J. Biomol. NMR. 2012; 54:881.

Figure 1.

Schematic view of the studied gp41¹⁻¹⁹⁴ construct, including the fusion peptide (FP), N-terminal heptad repeat (NHR), immuno-dominant loop region (IL), C-terminal heptad repeat (CHR), membrane proximal external region (MPER) and transmembrane helix (TM) anchoring gp41 to the viral envelope. The numbering 1–194 refers to the gp41¹⁻¹⁹⁴ construct; the numbering 512–705 refers to the location in the *Env* precursor.

Figure 2. 800 MHz ^{15}N - ^1H TROSY-HSQC spectrum of 0.5 mM ^2H , ^{15}N , ^{13}C gp41¹⁻¹⁹⁴ in DPC micelles, 50mM sodium acetate, pH 4.0, 25 mM KCl, recorded at 40 °C.

Figure 3.

^{15}N relaxation data recorded for gp41^{1–194} in DPC micelles: a) ^{15}N R_1 relaxation data recorded at 600 MHz (black) and 800 MHz (red) are highly consistent for both fields. b) $R_{2,0}$ relaxation data (derived from $R_{1\rho}$ with a 2 kHz RF field; R_1 contribution corrected), at 600 MHz (black) and 800 MHz. c) $^{15}\text{N}\{-^1\text{H}\}$ NOE values. d) Transverse CSA-dipolar cross-correlated relaxation rates η_{xy} . e) Hahn-echo transverse relaxation rates, $R_{2\beta}$, at 600 MHz (black) and 800 MHz (red) for the slowly relaxing component of the $^{15}\text{N}\{-^1\text{H}\}$ doublets. Unlike the $R_{1\rho}$ experiment, all conformational exchange effects, $R_{2\text{ex}}$, contribute to $R_{2\beta}$ and are not re-focused.

Figure 4.

Homotrimeric gp41 shows a high degree of intrinsic mobility on the nanosecond and microsecond time scale. NMR and SAXS data are compatible with the pre-hairpin intermediate sampling a range of relative CHR vs NHR orientations (A), possibly in exchange with a low population of the late-fusion 6HB (C). A conformational change from the pre-hairpin intermediate to the late-fusion 6HB requires breaking up of the CHR three-helical bundle with concomitant disorder of the CHR and MPER region during the transition period (B). The large exchange broadening for CHR and MPER region is indicative of such a scenario. No NMR information on the relative arrangement of CHR, MPER and TM domains is available and the arrangement shown is purely schematic. Intense resonances observed for FP exclude a strong interaction between FP and TM, although weak PRE interactions point to their transient proximity.

Table 1Results of the extended model-free analysis of ^{15}N relaxation data.

Cluster	τ_c [ns] ^[a]	S^2_f ^[b]	τ_s [ns] ^[c]	S^2_s ^[d]
I4-L12	44±3	0.77±0.02	5.4±0.7	0.42±0.04
Q29-R46	44±3	0.85±0.02	4.0±0.9	0.64±0.05
I62-L76	44±3	0.90±0.04	3.1±0.9	0.76±0.03
T18-M24	44±3	0.62±0.02	2.3±0.2	0.37±0.02
G86-N105	44±3	0.67±0.03	2.5±0.7	0.48±0.04

^[a] Rotational correlation time.^[b] Order parameter for the fast time scale, order parameters can range between 0 and 1 with 0 being fully mobile and 1 entirely rigid. The correlation time of the fast time scale motion has been fixed to 50 ps during the fitting.^[c] Correlation time of the slow time-scale internal motion ($\ll \tau_c$).^[d] Order parameter for the slow time scale.