

Complete dissociation of the HIV-1 gp41 ectodomain and membrane proximal regions upon phospholipid binding

Julien Roche · John M. Louis · Annie Aniana ·
Rodolfo Ghirlando · Ad Bax

Received: 16 November 2014 / Accepted: 17 January 2015 / Published online: 29 January 2015
© Springer Science+Business Media Dordrecht (outside the USA) 2015

Abstract The envelope glycoprotein gp41 mediates the process of membrane fusion that enables entry of the HIV-1 virus into the host cell. Strong lipid affinity of the ectodomain suggests that its heptad repeat regions play an active role in destabilizing membranes by directly binding to the lipid bilayers and thereby lowering the free-energy barrier for membrane fusion. In such a model, immediately following the shedding of gp120, the N-heptad and C-heptad helices dissociate and melt into the host cell and viral membranes, respectively, pulling the destabilized membranes into juxtaposition, ready for fusion. Post-fusion, reaching the final 6-helix bundle (6HB) conformation then involves competition between intermolecular interactions needed for formation of the symmetric 6HB trimer and the membrane affinity of gp41's ectodomain, including its membrane-proximal regions. Our solution NMR study of the structural and dynamic properties of three constructs containing the ectodomain of gp41 with and without its membrane-proximal regions suggests that these segments do not form inter-helical interactions until the very late steps of the fusion process. Interactions between the polar termini of the heptad regions, which are not associating with the lipid surface, therefore may

constitute the main driving force initiating formation of the final post-fusion states. The absence of significant intermolecular ectodomain interactions in the presence of dodecyl phosphocholine highlights the importance of trimerization of gp41's transmembrane helix to prevent complete dissociation of the trimer during the course of fusion.

Keywords Backbone dynamics · Chemical shift perturbation · Hemagglutinin · Membrane fusion · MPER · Pre-hairpin intermediate

Introduction

The first step of human immunodeficiency virus infection involves fusion of the viral and target cell membranes, a process mediated by its viral envelope glycoproteins, gp120 and gp41 (Harrison 2008; Blumenthal et al. 2012). These envelope proteins form a non-covalent C₃-symmetric homotrimeric complex on the surface of the virus, where three gp120 surface subunits sequester a trimer of gp41 proteins that are anchored to the viral surface by their transmembrane helix (TM), in a 3-helical bundle arrangement (Mao et al. 2013b). Binding of gp120 to the cell surface receptors CD4 and chemokine receptors CXCR4 or CCR5 triggers a cascade of conformational changes that disrupt the interactions between gp41 and gp120, and results in an extended gp41 conformation (Roux and Taylor 2007; Harrison 2008; Merk and Subramaniam 2013). In this extended pre-fusion state, the highly hydrophobic N-terminal fusion peptide (FP) of gp41 becomes exposed and thereby available to insert itself into the host cell membrane, while its TM keeps the protein anchored to the viral membrane (Furuta et al. 1998; Bartesaghi et al.

Electronic supplementary material The online version of this article (doi:10.1007/s10858-015-9900-4) contains supplementary material, which is available to authorized users.

J. Roche · J. M. Louis · A. Aniana · A. Bax (✉)
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA
e-mail: bax@nih.gov

R. Ghirlando
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

2013; Julien et al. 2013; Lyumkis et al. 2013). After the host cell and viral membranes have fused, the gp41 ectodomain which links the FP and TM regions, adopts a very different structure. Fully analogous to the extensively studied class I viral fusion protein hemagglutinin of the influenza virus (Skehel and Wiley 2000), in the post fusion state the ectodomain is found in a helical hairpin conformation and assembles as a symmetric 6-helix bundle (6HB) homotrimer (Chan et al. 1997; Caffrey et al. 1998; Markosyan et al. 2003), placing the FP in physical proximity to the TM domain (Fig. 1). The refolding of gp41 trimers into the highly stable 6HB arrangement is believed to overcome the large free-energy barrier of membrane fusion (Gallo et al. 2003; Blumenthal et al. 2012).

The current post-fusion structural models are mainly based on X-ray crystallographic studies of the soluble gp41 ectodomain, composed of N- and C-terminal heptad repeat (NHR and CHR) helical regions (Fig. 2), which show that the CHR helices pack in an antiparallel manner into the conserved hydrophobic grooves formed on the surface of the central trimeric bundle of parallel NHR helices (Chan et al. 1997; Tan et al. 1997; Weissenhorn et al. 1997; Caffrey et al. 1998; Buzon et al. 2010). On the pre-fusion side, the picture is less clear. It is generally accepted that upon CD4 binding, the trimeric gp120-gp41 complex transitions from its “resting” pre-fusion state to an activated pre-fusion state. However, large differences in the arrangements of the gp41 ectodomain have been reported for the resting state, where Mao et al. reported cryo-electron microscopy (cryo-EM) results which show a torus-like packing of the ectodomains, each broken into seven α -helices, whereas other cryo-EM (Bartasaghi et al. 2013; Lyumkis et al. 2013) and X-ray (Julien et al. 2013) studies find a very different arrangement in which the NHR and CHR helices remain largely intact. The cause of this discrepancy remains the subject of much debate (Henderson

2013; Mao et al. 2013a; Subramaniam 2013; van Heel 2013). However, consensus appears to exist on the structure of the activated state of the pre-fusion complex (Bartasaghi et al. 2013; Lyumkis et al. 2013) which shows a structural arrangement of the ectodomain that differs significantly from the post-fusion 6HB conformation. In the activated pre-fusion gp120-gp41 complex, the NHR regions form a central 3-helical bundle, but the CHR helices do not fold back along the NHR trimer in an antiparallel manner and instead extend away from the bottom of the trimer (Fig. 1a). No intermolecular interactions are seen for the CHR helices in the X-ray structure, but intramolecular contacts between the C-terminal end of the NHR helix and the middle of the CHR helix are believed to stabilize this arrangement (Julien et al. 2013).

This activated pre-fusion complex then sheds its gp120 domains, and the gp41 ectodomains are believed to transition to an extended pre-hairpin intermediate (PHI, Fig. 1b). No direct structural data are available that detail how the trimer subsequently switches from the PHI to the post-fusion 6HB state (Fig. 1e). Indeed, because of their intrinsic dynamical and multi-state nature, the intermediate states of the Env-driven fusion are not directly accessible to the standard X-ray crystallography or cryo-EM techniques, and therefore their structural features remain largely unknown (Tamm et al. 2014). However, the critical importance of understanding the structural details of these intermediate states for the design of next generation fusion inhibitor peptides is well recognized. In particular, the emergence of a new class of membrane-conjugated peptide fusion inhibitors suggests direct interaction of the heptad regions with the cellular and/or viral membranes during the course of the fusion process (Hildinger et al. 2001; Melikyan et al. 2006; Hollmann et al. 2013). Indeed, a growing body of evidence based on the activities of these NHR- and CHR-derived fusion-inhibiting peptides

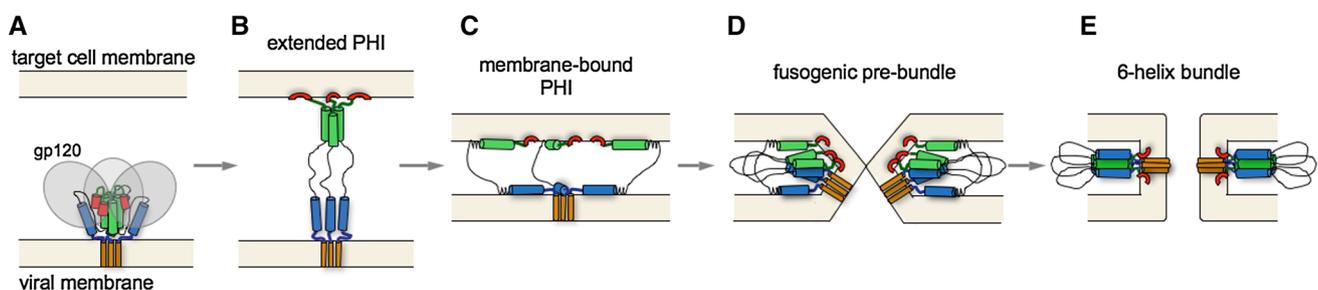


Fig. 1 Model of the intermediate steps in gp41-driven fusion of the viral and target cell membranes, showing the NHR (light green) and CHR (light blue) segments, the membrane proximal regions FPPR (dark green) and MPER (dark blue), and the membrane-anchoring elements, FP (red) and TM (orange). **a** The activated pre-fusion complex formed by three gp120 subunits (light gray) sequestering a trimer of gp41 at the surface of the viral membrane. **b** The short-lived

extended pre-hairpin intermediate state. **c** The collapsed PHI state, where NHR and CHR have become embedded in the viral and host-cell membranes, thereby pulling the membranes into juxtaposition. **d** The fusogenic pre-bundles, possibly initiated by contacts between the short polar segments at opposing ends of the NHR and CHR. **e** The post-fusion 6HB trimers, stabilized by FP-TM, FPPR-MPER, and 6HB NHR-CHR interactions

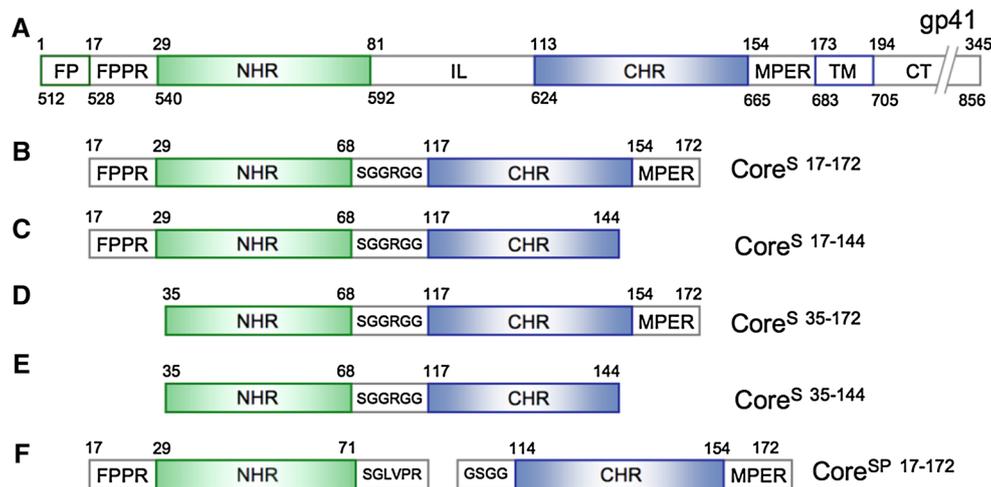


Fig. 2 **a** Schematic representation of the gp41 sequence, including the fusion peptide (FP), the fusion peptide proximal region (FPPR), the N-terminal heptad repeat (NHR), the immunodominant loop (IL), the C-terminal heptad repeat (CHR), the membrane proximal external region (MPER) and the trans-membrane domain (TM). The numbering 512–856 refers to the Env precursor sequence, while the 1–345 numbering corresponds to the gp41 sequence. For all constructs used in the present study, the IL loop was replaced by a 6-residue linker connecting the NHR and CHR regions. The new recombinant proteins

were designed as extensions of the previously described short ectodomain construct $\text{Core}^{\text{S}}_{35-144}$ (**e**) (named Core^{S} in Roche et al. 2014 and contains either the FPPR ($\text{Core}^{\text{S}}_{17-144}$) (**c**), the MPER ($\text{Core}^{\text{S}}_{35-172}$) (**d**), or both membrane proximal regions ($\text{Core}^{\text{S}}_{17-172}$) (**b**). (**F**) $\text{Core}^{\text{SP}}_{17-172}$, soluble at pH 6 in the absence of DPC, was generated by thrombin cleavage of a separate construct. Note that in $\text{Core}^{\text{SP}}_{17-172}$ the NHR and CHR helices each extend by three residues towards the IL, and the chains also contain several additional residues from the thrombin cleavage site

suggests that the gp41 ectodomain interacts directly with the membranes and actively participates in the fusion process (Kliger et al. 2000; Sackett and Shai 2002; Korazim et al. 2006; Lev et al. 2009). This lipid-binding property has been postulated to facilitate membrane fusion by introducing an additional destabilization of the viral and target cell membranes, thereby lowering the free-energy barrier for fusion (Korazim et al. 2006).

In two recent reports, Lakomek et al. (2013, 2014) carried out NMR studies in dodecyl phosphocholine (DPC) detergent of two gp41 constructs that both included the full ectodomain, the membrane proximal region, and the TM helix (residues 1–194; Fig. 2). In the first gp41 study (Lakomek et al. 2013), which also included the fusion peptide (FP) and fusion peptide proximal region (FPPR), nearly all of the FP, FPPR and most of the NHR and inter-helical linker (IL) regions gave rise to relatively sharp resonances. Chemical shifts for the FP region agreed closely with those studied earlier for the free fusion peptide, known to adopt a largely α -helical conformation based on NOE and residual dipolar coupling measurements (Jaroniec et al. 2005). FPPR residues were found to have slower transverse ^{15}N relaxation rates and low heteronuclear ^{15}N - $\{^1\text{H}\}$ NOE values which, combined with $^{13}\text{C}^{\alpha}$ secondary chemical shifts that were intermediate between random coil and α -helical values, pointed to a high degree of dynamic disorder for this linker, while transiently adopting α -helical conformations. The region extending from T27 to L76,

encompassing the NHR helix as well as a number of residues immediately preceding and following this helix, showed rather uniform ^{15}N relaxation behavior and $^{13}\text{C}^{\alpha}$ chemical shifts fully consistent with a mostly α -helical conformation. However, the amide signals of residues I49–V59 were unobservable, presumably as a result of intermediate rate conformational exchange (Lakomek et al. 2013). Most of the IL residues, from L81 to L108, showed relatively homogeneous, elevated fast internal dynamic behavior, rather similar to the FPPR region. Reported $^{13}\text{C}^{\alpha}$ chemical shifts for IL again were intermediate between random coil and α -helical values, but also included a stretch (T95–S104) with near random coil values (Lakomek et al. 2013). Very few of the CHR, MPER and TM resonances could be observed, a finding attributed to conformational exchange on an intermediate time scale. The subsequent study of a shorter construct, lacking the FP and N-terminal half of FPPR, yielded much better NMR spectra, but chemical shifts and ^{15}N dynamics that were virtually unchanged relative to the longer construct (Lakomek et al. 2014). Based on analytical ultracentrifuge measurements and small angle X-ray scattering data, both samples were reported to be fully homotrimeric at the concentrations ($>100\ \mu\text{M}$) used in these NMR studies. ^{15}N relaxation data, when fitted with the extended Lipari-Szabo model-free method (Lipari and Szabo 1982; Clore et al. 1990), indicated an overall tumbling time of 44 ns, and substantial amplitude internal motions on the time scale of

a few ns. Interestingly, the shorter construct also yielded resonances for the N-terminal half of the CHR helix, up to residue L134, showing α -helical $^{13}\text{C}^\alpha$ chemical shifts and ^{15}N relaxation rates very similar to those seen for the NHR helix. However, the vast majority of the C-terminal end of CHR, MPER, and all but the last five residues of TM remained invisible, indicative of conformational exchange. Although the IL region was found to increase trimerization affinity, the NHR region was concluded to be principally responsible for trimerization (Lakomek et al. 2014).

In our recent study of a much shorter ectodomain construct (Fig. 2E), addition of either DPC or small unilamellar vesicles, of a lipid composition that mimics the T cell membrane, resulted in the breakup of the trimeric 6HB state of this domain, yielding flexibly linked NHR and CHR monomeric α -helical segments, located at the water–lipid interface. A very similar arrangement of the NHR and CHR helices was found in the presence of lipids when linking these helices by the full-length IL linker, rather than the 6-residue linker (Roche et al. 2014). Based on these data, we proposed a new model for the fusion pathway where, after the gp120 segments have been shed, the CHR and NHR helices first embed in the viral and host cell membranes, respectively, then permitting the transition to 6HB formation. By melting into the two membranes, the ectodomain then provides the force for pulling these into close proximity, limited by the length of the IL region which is also lipophilic (Lakomek et al. 2013). In this model (Fig. 1), the subsequent zippering to the final 6HB trimeric form of the ectodomain in the late stage of the fusion process counterbalances its lipid binding affinity (Roche et al. 2014). This model is reminiscent of the splay-melting model, first proposed for hemagglutinin on the basis of EPR data (Yu et al. 1994; Carr and Kim 1994). The model was subsequently criticized on the basis of crystallographic data and a photolabeling study (Chen et al. 1995; Durrer et al. 1996), but strongly supported by vastly increased fusogenicity when extending the hemagglutinin fusion peptide by its ectodomain (Epand et al. 1999).

In our fusion model (Fig. 1) the NHR helices (as well as the CHR helices) fully dissociate and “melt” into the membrane surface as soon as the gp120 units are shed and FP engages the host cell membrane. This contrasts with most other models where the NHR helices serve as the “glue” for keeping the protein trimeric (Harrison 2008; Lakomek et al. 2014; Tamm et al. 2014; Sackett et al. 2014). A potential concern in evaluating the NHR dissociation, observed for our truncated Core^S segment (Fig. 2e) upon interaction with lipids, relates to the fact that Core^S lacks the other ectodomain segments which could impact the trimerization propensity (Banerjee and Weliky 2014). In particular, a recent NMR study of MPER in DPC, where this fragment was forced into a homotrimeric assembly by

generating it as a C-terminal extension of the homotrimeric foldon domain of T4 fibrin, showed specific intermolecular interactions between these MPER peptides (Reardon et al. 2014).

In another recent crystallographic study of the trimeric gp41 ectodomain that included the FPPR and MPER segments, a 6HB state quite similar to that of the shorter Core^S was observed, but also showing the existence of hydrophobic and hydrogen-bond interactions between the FPPR and MPER segments (Buzon et al. 2010). Although in all current fusion models such interactions cannot form until the opposing ends of the ectodomain have come into close spatial proximity, we previously speculated that such interactions might be responsible for refolding the monomeric, membrane-associated ectodomains into the final 6HB state (Roche et al. 2014). Whether such FPPR–MPER interactions are transiently formed in the pre-hairpin states, and indeed represent a driving force for the formation of the final 6HB, is key to understanding the viral fusion mechanism.

In our present study, we compare the solution behavior of the Core^S construct used in our earlier study with ectodomain constructs that now also include FPPR, MPER, or both regions. Our results show that the helical segments in the presence of DPC do not significantly interact with one another, and that all of these are monomeric. By comparison with the results for the full length gp41 domain, which additionally includes the TM helix (Lakomek et al. 2013), this observation strongly suggests that the TM element of gp41 is responsible for maintaining the trimeric state of the protein after shedding of gp120.

Materials and methods

Three gp41 constructs, Core^S_{17–172}, Core^S_{17–144} and Core^S_{35–172} were synthesized and cloned into the PJ414 vector between the NdeI and BamHI sites, for expression in BL21 (DE3) host (DNA 2.0, Menlo Park, CA). Typically, cells harvested from 0.5 L of culture were suspended in 70 mL of buffer A [50 mM Tris–HCl, pH 8, 10 mM ethylenediaminetetraacetic acid (EDTA) and 10 mM dithiothreitol (DTT)], followed by the addition of lysozyme (100 $\mu\text{g mL}^{-1}$) for cell lysis and sonication at 4 °C. The insoluble recombinant protein was washed by resuspension in 70 mL of buffer B (50 mM Tris–HCl, pH 8, 10 mM EDTA, 10 mM DTT, 2 M urea, and 1 % Triton X-100) and subsequently in buffer A. In all cases, the insoluble fraction was pelleted by centrifugation at 20,000 g for 30 min at 4 °C. The final pellet was solubilized in 50 mM Tris–HCl, pH 8.0, 7.5 M guanidine hydrochloride (GnHCl), 5 mM EDTA, and 10 mM DTT (buffer SB) to yield a concentration of 20 mg mL⁻¹. Buffer SB also

contained 10 mM DPC to facilitate complete solubility of Core^S 17–144 and Core^S 35–172. Samples were spun at 12,800 rpm in an Eppendorf centrifuge for 45 min and supernatants were recovered. A maximum of 18 mg of protein was applied on a Superdex-75 column (HiLoad 1.6 cm, 60 cm, GE HealthCare, Piscataway, NJ) equilibrated in 50 mM Tris–HCl, pH 8, 4 M GnHCl, 5 mM EDTA, 1 mM DTT and 2 mM DPC at a flow rate of 1.5 mL min⁻¹ at ambient temperature. Peak fractions were pooled and subjected to reverse-phase HPLC on POROS 20 R2 resin (PerSeptive Biosystems, Framington, MA) and eluted using a linear gradient from 99.95 % water (v/v) and 0.05 % TFA to 60 % acetonitrile (v/v), 0.05 % TFA (v/v), and 39.95 % water (v/v) over a period of 16 min at a flow rate of 4 mL min⁻¹. Peak fractions were combined, estimated for protein content at 280 nm, and stored at –70 °C. The various constructs were verified both by DNA sequencing and electrospray ionization mass spectrometry (ESI–MS). In the case of Core^S 17–144, resuspension in buffer B which includes 2 M urea was avoided because of decreased sample insolubility resulting in poorer yield from the insoluble fraction (pellet) during the purification.

Another construct, termed Core^{SP} 17–172, was cloned, expressed and purified from inclusion bodies, similar to the protocol described for the other constructs. A thrombin site was engineered within the L6 spacer. Thus, regions spanning residues 17–71 and 114–172 are separated by the spacer, SGLVPRGSGG. Following purification of the solubilized inclusion bodies by size exclusion chromatography on Superdex-75, peak fractions corresponding to the full length protein were pooled and dialyzed extensively against 25 mM Tris–HCl, pH 7.5, 150 mM NaCl and 2 mM CaCl₂. Thrombin digestion was carried out for ~1 day at ambient temperature in the presence of 5 mM DPC to facilitate solubility of the full-length protein. Following nearly complete cleavage as monitored by SDS-PAGE, the reaction mixture was concentrated, adjusted to a final concentration of 6 M GnHCl and loaded onto the Superdex-75 column equilibrated in 4 M GnHCl, 20 mM sodium formate buffer, pH 2.6. Peak fractions of fragments 17–71 (6.4 kDa) and 114–172 (7.4 kDa), which co-elute later than the miniscule amount of undigested full-length protein, were combined and dialyzed according to the scheme described.

Isotope enriched proteins were grown in minimal media with the appropriate isotope source in 99 % D₂O. An aliquot of the protein was dialyzed in 25 mM sodium formate at pH 3 to exchange acetonitrile, followed by extensive dialysis in 50 mM sodium acetate at pH 5 and it was subsequently concentrated. For the constructs which encompass the FPPR and/or the MPER segments, DPC at a final concentration of 5 mM was added prior to dialysis in 25 mM sodium formate at pH 3. NMR samples with uniform ca 99 % ²H, ¹³C, ¹⁵N enrichment were prepared to

final concentrations of 0.4–0.7 mM in ca 100 mM perdeuterated DPC (Anatrace), 93 % H₂O/7 % D₂O (v/v).

¹H–¹⁵N TROSY-HSQC spectra were recorded on a uniformly (>95 %) ²H/¹⁵N/¹³C-enriched sample at 0.4 mM (monomer concentration) in 50 mM sodium acetate pH 4.0 or 25 mM sodium phosphate pH 6.0 at 310 K, in the presence of 100 mM DPC, using a 600 MHz Bruker Avance II spectrometer, equipped with a z-axis TCI cryogenic probe. Backbone assignments were based on 3D TROSY-HNCO and 3D TROSY-HNCACB spectra recorded at 600 MHz on the ²H/¹⁵N/¹³C-enriched sample at 0.4 mM in 50 mM sodium acetate pH 4.0, at 310 K, and confirmed by 3D ¹⁵N-separated NOESY-HSQC spectra.

The ¹⁵N R₁, R₂ and steady-state heteronuclear ¹⁵N–{¹H} NOE data were collected at a ¹H frequency of 600 MHz using TROSY-based ¹H–¹⁵N heteronuclear experiments (Lakomek et al. 2012) on uniformly ²H/¹⁵N-labeled samples in the presence of 100 mM of DPC. The R₂ rates were derived from R_{1ρ} values measured with a radiofrequency 1.3 kHz spin-lock field by correcting them in the standard manner for ¹⁵N radiofrequency offset (Cavanagh et al. 2007). All relaxation experiments were performed at a protein concentration of 0.4 mM in 50 mM sodium acetate pH 4.0, at 310 K.

Results

The earlier NMR study by Lakomek et al. (2013) of the 194-residue construct gp41^{1–194} in DPC micelles indicated that the general behavior of the protein was very similar at pH 4.0 and 7.1, but that at the elevated temperature (40 °C) used for the NMR study the lower pH was preferred, in part due to slower hydrogen exchange of amide protons in flexible regions of the protein. In order to allow a direct comparison with chemical shifts reported in that study, and in our previous study of Core^S, data were again recorded at pH 4.

In our previous study of the gp41 ectodomain, we found that in the presence of DPC the IL segment extends the NHR helix in the C-terminal direction by two turns of helix, and similarly, that the C-terminal end of IL also adopts an α-helical conformation, connected by a kink to the CHR helix. Even though the IL region has been reported to stabilize trimerization of the NHR (Lakomek et al. 2014), substituting this region by the short SGGRGG linker in our experiments had no impact on the trimer–monomer transition, and did not impact the chemical shifts (and thereby the structure or interaction mode) of the NHR and CHR helices in the presence of detergent (Roche et al. 2014). Therefore, when evaluating the effects of extending the NHR and CHR helices into the N- and C-terminal directions, respectively, by extending the construct with the

FPPR and MPER segments, we again used the SGGRGG linker segment, yielding less congested NMR spectra.

Here we studied three constructs—Core^{S17–172}, Core^{S35–172} and Core^{S17–144}—which complement our previous study of Core^S and Core^{IL} (Fig. 2). These constructs are therefore strict extensions of our previous truncated ectodomain construct (named Core^{S35–144} in the new nomenclature, Fig. 2e).

Thermal stability and oligomeric state of the ectodomain constructs

Circular dichroism spectra of Core^{S17–172} recorded at pH 4.0 in the absence of DPC show the characteristic signature of an α -helical protein with a minimum at 222 nm, corresponding to ca 95 % helical content (Fig. S1A). In contrast to our previous shorter ectodomain construct (Core^{S35–144}), the addition of 10 mM DPC to Core^{S17–172} did not induce any substantial change in helicity (Fig. S1A, B). Nevertheless, clear differences in thermal stability appear between the DPC-bound and DPC-free states of Core^{S17–172}. While both the shorter Core^{S35–144} and the longer Core^{S17–172} show a midpoint of thermal denaturation of ca 79 °C in the absence of DPC, both proteins exhibit a characteristic non-cooperative decrease in ellipticity with temperature in the presence of DPC, with melting temperatures outside the range accessible in our experiments (Fig. S1C, D). Similar thermostable behavior has been reported previously for other ectodomain constructs in the presence of detergent (Lev et al. 2009; Sackett et al. 2009).

The molar masses of the species populated in the absence and in the presence of DPC were directly measured using sedimentation velocity (SV) experiments. In the absence of detergent, a narrow SV c(s) distribution corresponding to a single population of trimers was observed for both Core^{S35–144} and Core^{IL35–144}, while a minor population with an average molar mass corresponding to a tetramer was observed for Core^{S17–172} in addition to the major trimer species (Fig. S2). Higher-order complexes have been reported recently for several ectodomain constructs in the absence of detergent (Gao et al. 2013; Banerjee and Weliky 2014) and may reflect non-specific inter-molecular interactions between the membrane proximal regions. In the presence of DPC, the SV experiments for all three ectodomain constructs show the presence of a single species, corresponding to a protein monomer bound to a DPC micelle (Fig. S2 and Table S1).

Absence of long-range interactions from chemical shift analysis

Chemical shifts are exquisitely sensitive to even transient, weak interactions, which can be notoriously difficult to detect using standard ¹H–¹H NOE-based approaches. To

investigate the presence of potential interactions between the FPPR and MPER regions, we compared the chemical shifts measured for Core^{S17–172}, containing both membrane proximal regions, with those recorded for the recombinant proteins containing either FPPR (Core^{S17–144}) or MPER (Core^{S35–172}) (Fig. 2 and Table S2). To probe the presence or absence of inter-helical interactions, we focused primarily on the backbone ¹³C α chemical shifts as widely used reporters on α -helicity, and on the H^N chemical shifts which are particularly sensitive to intermolecular and long-range interactions.

The chemical shifts measured for the Core^{S17–144} construct containing the FPPR region show a very close correspondence with those measured under the same conditions for the longer construct Core^{S17–172}, containing both FPPR and MPER regions (Fig. 3a–c). Except for the large chemical shift differences seen for the C-terminal residues of Core^{S35–144}, reflecting the difference in covalent structure, the largest difference in ¹³C α chemical shift is 0.14 ppm (I131) and for H^N it is 0.07 ppm (A50). Excluding the 6 C-terminal residues, the overall root mean square difference (rmsd) between the chemical shifts measured for these two recombinant proteins is only 0.05 and 0.02 ppm for the ¹³C α and H^N chemical shifts, respectively. Similarly, we observed a very close correspondence between the backbone chemical shifts measured for Core^{S35–172} and Core^{S17–172} (Fig. 3d, f). Excluding the 10 N-terminal residues of Core^{S35–172}, which again reflect the sequence differences between these two constructs, an rmsd of 0.06 and 0.02 ppm is observed when comparing the secondary ¹³C α and H^N chemical shifts, respectively. If the FPPR region were to make significant transient interactions with MPER on the surface of the micelle, chemical shift perturbation (CSP) for residues 17–29 of Core^{S17–144} relatively to Core^{S17–172} would be expected. Similarly, transient FPPR–MPER would impact the chemical shifts of residues 154–172 of Core^{S35–172} relatively to Core^{S17–172}.

A similar chemical shift comparison between our three ectodomain constructs was also performed at pH 6.0. Since Core^{S17–172} is poorly soluble at pH 6.0, we generated an alternative construct similar to the one used by Weissenhorn and co-workers (Buzon et al. 2010), named Core^{SP17–172}, prepared as a 1:1 mixture of FPPR-NHR and CHR-MPER fragments (see Materials and Methods and Fig. 2f). Besides the terminal regions reflecting the sequence differences between the constructs, we observed again a very close correspondence between the amide proton chemical shifts measured at pH 6.0 for Core^{SP17–172} and those measured under the same conditions for Core^{S17–144} (Fig. S3A) and Core^{S35–172} (Fig. S3B), with an rmsd of 0.03 and 0.02 ppm for the ¹³C α and H^N chemical shifts, respectively.

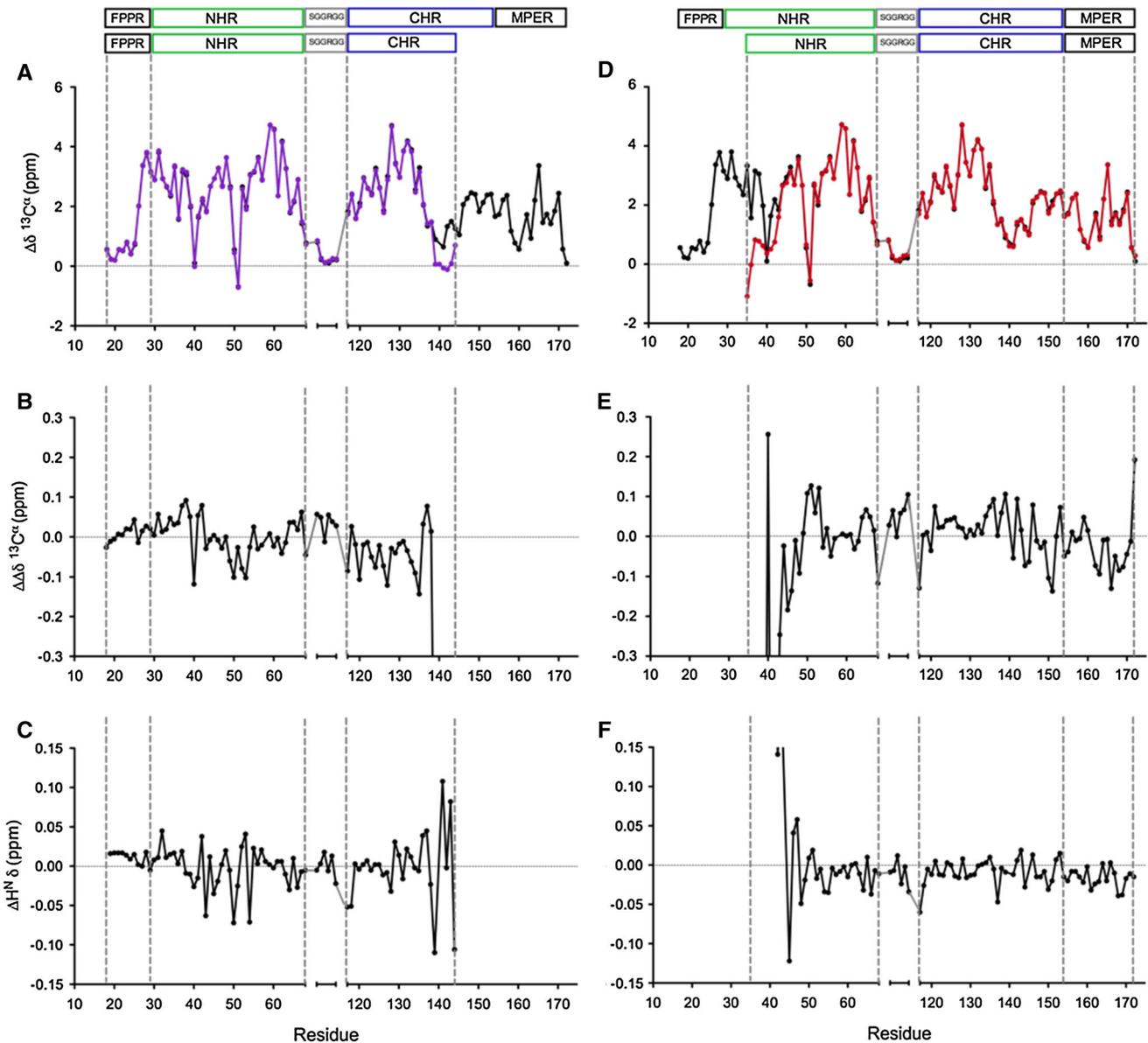


Fig. 3 Comparison of the backbone chemical shifts recorded for $\text{Core}^{\text{S}}_{17-144}$, $\text{Core}^{\text{S}}_{35-172}$ and $\text{Core}^{\text{S}}_{17-172}$ in 50 mM sodium acetate (pH 4.0) in the presence of 100 mM DPC. The length of each region (FPPR, NHR, CHR and MPER), is depicted on the top of each panel. **a, d** Secondary backbone $^{13}\text{C}^\alpha$ chemical shifts recorded for Core^{S}

$_{17-144}$ (purple), $\text{Core}^{\text{S}}_{35-172}$ (red) and $\text{Core}^{\text{S}}_{17-172}$ (black). **b, c** Chemical shifts difference between $\text{Core}^{\text{S}}_{17-144}$ and $\text{Core}^{\text{S}}_{17-172}$ for $^{13}\text{C}^\alpha$ and $^1\text{H}^{\text{N}}$. **e, f** $^{13}\text{C}^\alpha$ and $^1\text{H}^{\text{N}}$ chemical shift differences between $\text{Core}^{\text{S}}_{35-172}$ and $\text{Core}^{\text{S}}_{17-172}$

The very close correspondence of the chemical shifts observed here for these recombinant proteins at both pH 4.0 and pH 6.0 strongly points to the absence of significant interactions between the FPPR and MPER regions on the surface of the micelles (Fig. 3). The minor chemical shift changes that are observed appear consistent with very weak, transient interactions, a necessary consequence when locating multiple, flexibly linked helical segments of a protein on the surface of a small DPC micelle.

Comparison of chemical shifts to trimeric gp41²⁷⁻¹⁹⁴

Lakomek et al. (2014) reported nearly complete backbone chemical shifts for the NHR, IL, and the N-terminal half of the CHR of gp41²⁷⁻¹⁹⁴ in the presence of DPC. Since this recombinant protein was shown to be trimeric in the presence of DPC, we here compare the backbone chemical shifts measured for $\text{Core}^{\text{S}}_{17-172}$ with those reported by Lakomek et al. The differences between the two constructs

relate to the absence of the IL and TM domains in Core^S_{17–172}, as well as a 10-residue longer FPPR region. A direct comparison of chemical shifts is indeed possible because they were measured under very similar experimental conditions (i.e. pH 4.0, large excess of DPC, 310 K in the present study, 313 K for Lakomek et al. 2014). With the exclusion of the N- and C-termini and the residues at the IL junction (replaced by a 6-residue linker in our constructs), the chemical shifts measured for Core^S_{17–172} agree remarkably closely with those reported for gp41^{27–194} (Fig. 4a, b), showing an rmsd of 0.16 ppm ($R^2 = 97\%$) and 0.09 ppm ($R^2 = 93\%$) for the ¹³C^α and ¹H^N chemical shifts, respectively. This agreement is comparable to that between gp41^{27–194} and gp41^{1–194} (Lakomek et al. 2014). Although the IL loop has been reported to stabilize the trimeric form of gp41 (Lakomek et al. 2014), our previous study (Roche et al. 2014) showed that Core^S_{35–144} remains fully monomeric when the SGGRRG linker is replaced by IL (Core^{IL}_{35–144}) (see also Fig. S2 and Table S1). The very close correspondence between the IL chemical shifts observed in monomeric Core^{IL}_{35–144} and homotrimeric gp41^{27–194} (Fig. 4a, b) therefore argue strongly against specific intermolecular IL interactions.

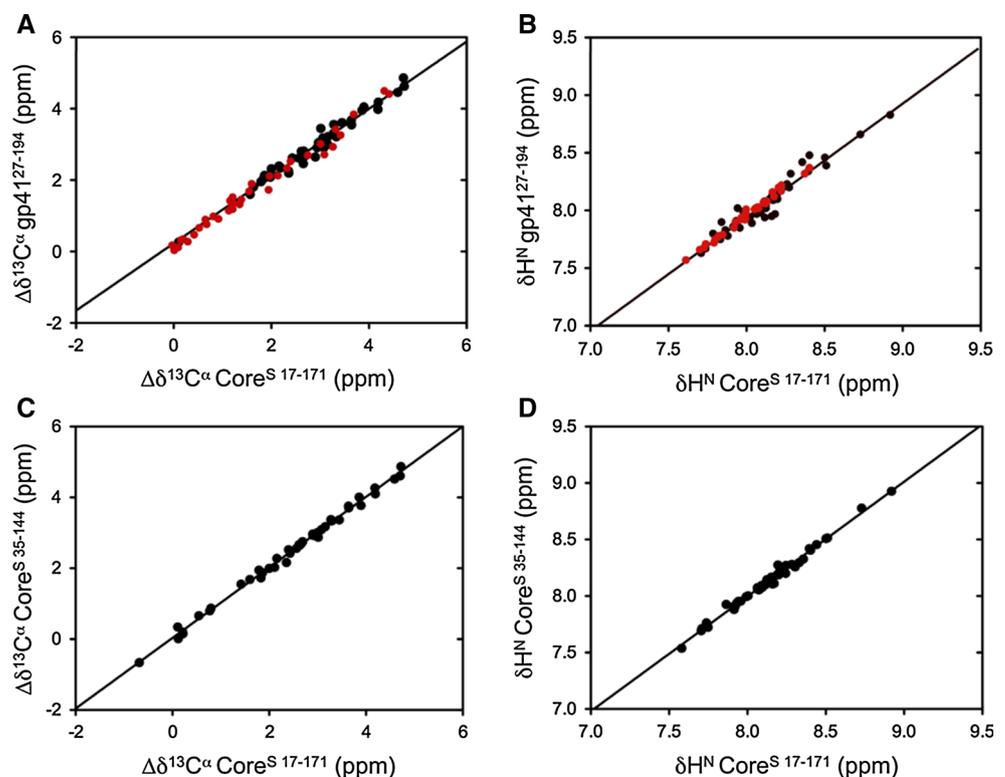
It is also interesting to compare the chemical shifts of the much smaller Core^S_{35–144}, lacking both the FPPR and MPER regions, with Core^S_{17–172}. In the presence of a large excess of DPC, Core^S_{35–144} was shown by size exclusion

chromatography coupled with multi-angle light scattering (SEC-MALS) measurements to be fully monomeric (Roche et al. 2014). Again, a very close correlation is observed between the chemical shifts of Core^S_{17–172} and Core^S_{35–144} (Fig. 4c, d), with an rmsd of 0.09 ppm ($R^2 = 99\%$) and 0.03 ppm ($R^2 = 99\%$) for the ¹³C^α and ¹H^N chemical shifts, respectively. This data argues strongly against any significant intermolecular interactions for Core^S_{17–172}. By inference, the very similar chemical shifts for Core^S_{17–172} and gp41^{27–194} point to the absence of intermolecular NHR interactions in gp41^{27–194}.

Comparison of ¹⁵N relaxation

Measurements of the ¹⁵N spin–lattice (R_1) and spin–spin (R_2) relaxation rates, together with the heteronuclear ¹⁵N–{¹H} NOE, were carried out for all three extended Core^S constructs: Core^S_{17–172}, Core^S_{17–144} and Core^S_{35–172} (Table S3). The ¹⁵N–{¹H} NOE is primarily sensitive to motions in the frequency range of ω_H ($2\pi \times 600$ MHz, for our measurements) and is often used as a proxy for fast internal motions. The profiles of observed ¹⁵N–{¹H} NOE values are remarkably similar for all three constructs, and also fall very close to those reported previously for the shorter Core^S_{35–144} construct (Fig. 5a, b). This data therefore indicates that the internal dynamics of these constructs is not impacted by distant extensions. By

Fig. 4 Comparison of chemical shifts. **a** Between the secondary ¹³C^α chemical shifts and **b** ¹H^N chemical shifts reported for the gp41^{27–194} recombinant protein (Lakomek et al. 2014) and those measured in the present study for Core^S_{17–172}. Chemical shifts measured in our previous study for the IL loop residues of the Core^{IL} construct (Roche et al. 2014) and those reported for gp41^{27–194} are shown in red. **c** Comparison between secondary ¹³C^α chemical shifts of Core^S_{35–144} (Roche et al. 2014) and Core^S_{17–172} (**d**) Comparison between ¹H^N chemical shifts of Core^S_{35–144} and Core^S_{17–172}



contrast, the ca eight residues closest to the N- and C-termini of each construct show depressed NOE values, indicative of increased internal dynamics, as expected on the basis of the reduced α -helical propensity (judged by $\Delta\delta^{13}\text{C}^\alpha$, Fig. 3a). Below average NOE values (ca 0.55) are also observed for residues 41–43 in the NHR region, immediately following Q40, a polar residue responsible for introducing a significant kink in the amphiphilic NHR helix (Roche et al. 2014). Interestingly, low NOE values (0.38–0.55) are also observed for the highly polar region

extending from residue 139–147, containing four Gln, two Asp, two Glu and a Lys residue. The increased internal dynamics observed for these residues likely reflects the inability of the polar side chains to engage the micelle surface in a contiguous α -helical conformation. Correspondingly, $^{13}\text{C}^\alpha$ secondary shifts for these residues, on average, fall about midway in between α -helical and random coil values (Fig. 3a). The very close agreement in both chemical shifts and heteronuclear $^{15}\text{N}\{-^1\text{H}\}$ NOE values measured for Core^S_{35–144}, Core^S_{17–144}, Core^S_{35–172}

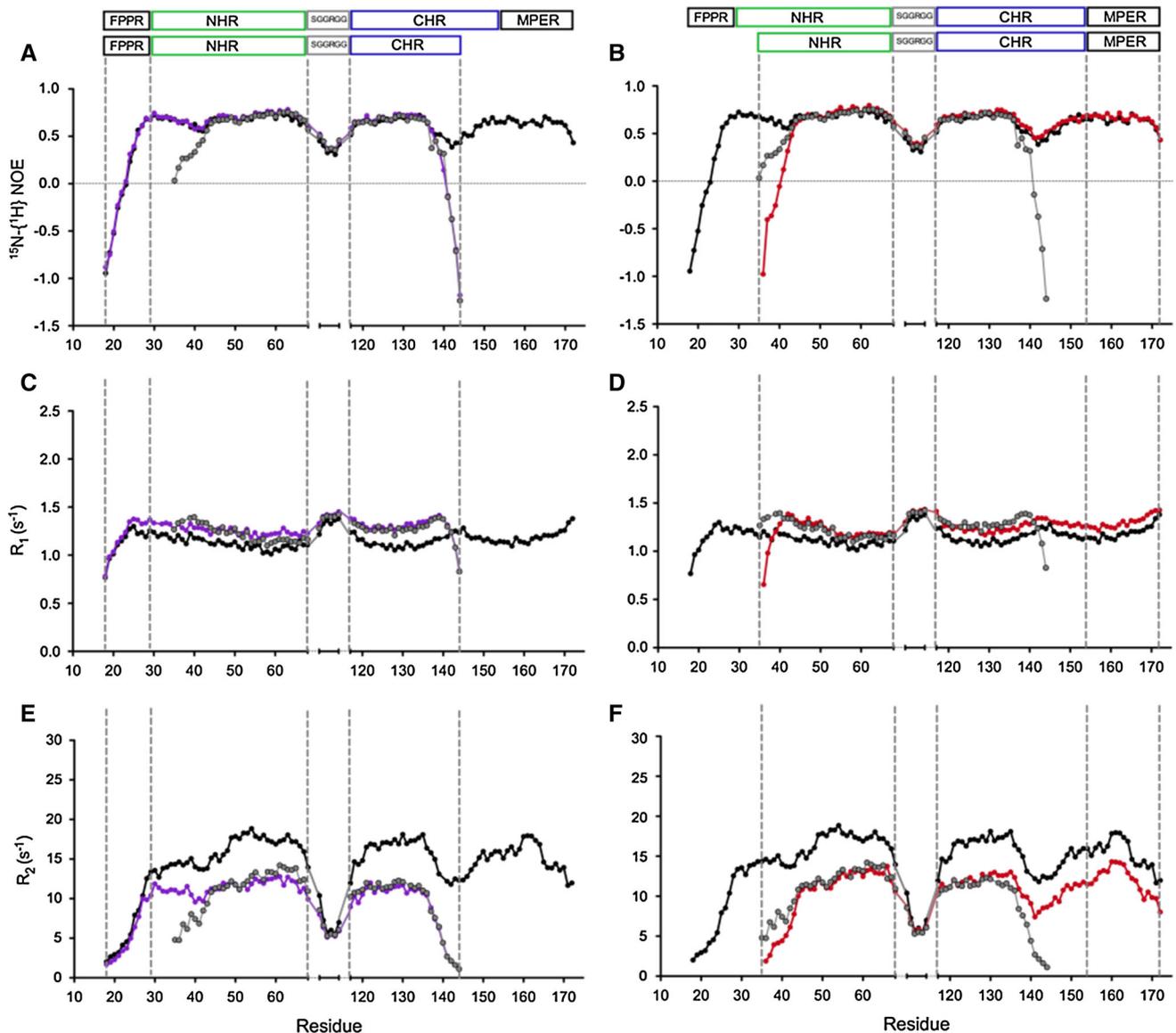


Fig. 5 Characterization of the internal dynamics of Core^S_{17–144}, Core^S_{35–172} and Core^S_{17–172} through ^{15}N relaxation measurements performed at a ^1H frequency of 600 MHz. The location of each region (FPPR, NHR, CHR and MPER) is depicted at the top of the figure. The comparison between Core^S_{17–144} (purple) and Core^S_{35–172} (black) is shown in the left half of the figure, with **a** the steady-state heteronuclear $^{15}\text{N}\{-^1\text{H}\}$ NOE values, **b** the R_1 relaxation rates and

c the R_2 relaxation rates. Similarly, the comparison between Core^S_{35–172} (red) and Core^S_{35–144} (black) is displayed in the right half of the figure, with **b** the $^{15}\text{N}\{-^1\text{H}\}$ NOE values, **d** the R_1 , and **(f)** the R_2 relaxation rates. For comparison, the heteronuclear $^{15}\text{N}\{-^1\text{H}\}$ NOE, R_1 and R_2 values measured for Core^S_{35–144} (Roche et al. 2014) are also shown (grey dots)

and Core^S_{17–172} provide strong evidence against any stable long-range interactions in Core^S_{17–172}.

In contrast to the heteronuclear ¹⁵N-¹H NOE values, the R₂ and R₁ relaxation rates measured for Core^S_{17–144} and Core^S_{35–172} are systematically offset relative to those of the larger Core^S_{17–172} (Fig. 5c–f). The substantial increase in the R₂ relaxation rates (Fig. 5c, d) together with the decreased R₁ values (Fig. 5e, f) observed in the structured region of the ectodomain, indicate that Core^S_{17–172} tumbles more slowly than Core^S_{17–144} and Core^S_{35–172}, as expected based on its increased molecular mass. Even short of a full relaxation analysis, inspection of this figure reveals that the fractional change in relaxation rate with size of the protein is substantially larger for R₂ than for R₁. This result contrasts with the comparable fractional changes expected for a well-structured protein far in the slow tumbling limit, and therefore points to substantial contributions to R₁ from internal motions on an intermediate time scale. However, whether these dynamics correspond to “whole-body rocking motions” of a well-structured protein relative to the rather large detergent micelle, as observed for the hemagglutinin fusion peptide (Lorieau et al. 2011), or independent motions of the flexibly linked helical domains, requires a more quantitative analysis.

The ¹⁵N relaxation data recorded for the three recombinant proteins therefore were analyzed by means of the extended Lipari-Szabo model-free formalism (Lipari and Szabo 1982; Clore et al. 1990), yielding an overall tumbling time of 12.9 ± 0.8 ns for Core^S_{17–172}, 10.8 ± 0.7 ns for Core^S_{17–144} and 10.9 ± 0.7 ns for Core^S_{35–144}, compared with a tumbling time of 10.6 ns for Core^S_{35–144} (Table 1). These results compare to overall tumbling times of 41 ns for trimeric gp41^{27–194} (Lakomek et al. 2014) and 44 ns for trimeric gp41^{1–194} (Lakomek et al. 2013). The ca 3.2-fold faster overall tumbling rate of Core^S_{17–172} than for gp^{27–194} is inconsistent with the ca 33 % smaller mass of Core^S_{17–172}, and points to a different oligomeric state of the two proteins. Interestingly, the magnitudes and order parameters for the slow internal motions observed for the three constructs we studied are all very similar to one

another (Table 1) and comparable to the values found by Lakomek for trimeric gp41^{1–194} (Lakomek et al. 2013).

Interestingly, extending Core^S_{35–144} by either the FPPR or MPER region has very little effect on its ¹⁵N relaxation properties (Fig. 5), but extending the protein by both FPPR and MPER causes a significant increase in R₂ but not in R₁. The latter is dominated by the relatively fast rocking of the helical elements on the surface of the DPC micelle, whereas R₂ would be highly sensitive to transient switching to a lowly populated, slower tumbling state. The latter could potentially represent the homotrimeric state, as a very low population of such a state (≤ ~10 %) would not be expected to result in chemical shift changes much larger than observed in our study. Alternatively, the slower tumbling state could result if the protein is bound to two DPC micelles. Indeed, it appears plausible that it becomes difficult to accommodate four substantial size helical elements, covering an area of ca 1,200 Å², on the surface of a single DPC micelle. In support of the latter hypothesis, when lowering the protein concentration by 15 % while increasing the DPC concentration by 60 %, we observe a further increase in R₂ (Fig. 6), consistent with what would be expected for an equilibrium between proteins partitioned on a single or two micelles. If transient trimer formation were responsible for the increased R₂ of Core^S_{17–172}, the equilibrium should be shifted back to monomer at the lower protein/higher DPC conditions, inconsistent with the further R₂ increase under the diluted conditions.

Discussion

Despite intense research over the past several decades, the structural features of the fusion intermediates connecting the pre-fusion gp120-gp41 complex (Julien et al. 2013; Lyumkis et al. 2013) and the post-fusion 6HB states (Chan et al. 1997; Tan et al. 1997; Weissenhorn et al. 1997) remain largely unknown (Tamm et al. 2014). In a recent study, we have shown that the 6HB trimeric structure of a recombinant ectodomain, lacking the membrane-interacting domains, FP, FPPR, MPER and TM, dissociates into stable monomers upon binding to zwitterionic detergent micelles, and behaves analogously in the presence of vesicles that mimic the T cell membrane composition (Roche et al. 2014). Although several studies have pointed out the critical role played by electrostatic interactions in the fusion mechanism (Lev et al. 2009; Sackett et al. 2011; Ratnayake et al. 2015), we found that dissociation of the soluble trimer into monomers upon binding to zwitterionic micelles occurs at both pH 4.0 and pH 6.0 (Roche et al. 2014 and the present study). We also observed that the short ectodomain construct Core^S_{35–144} binds to the negatively charged T-cell mimicking vesicles (LM-3) at both

Table 1 Results of the extended model-free analysis of the ¹⁵N relaxation data

Construct	τ _c (ns)	S _f ²	τ _s (ns)	S _s ²
Core ^S _{17–172}	12.9 ± 0.8	0.91 ± 0.04	2.5 ± 0.8	0.80 ± 0.05
Core ^S _{17–144}	10.8 ± 0.7	0.93 ± 0.03	2.9 ± 0.6	0.82 ± 0.04
Core ^S _{35–172}	10.9 ± 0.7	0.93 ± 0.03	2.8 ± 0.6	0.81 ± 0.04
Core ^S _{35–144}	10.6 ± 0.6	0.92 ± 0.04	2.9 ± 0.6	0.83 ± 0.04

For all three constructs, the ¹⁵N relaxation data were fitted over the most highly ordered helical segments of the NHR (residues 49–67) and CHR (residues 121–138). All data collected at 600 MHz ¹H frequency

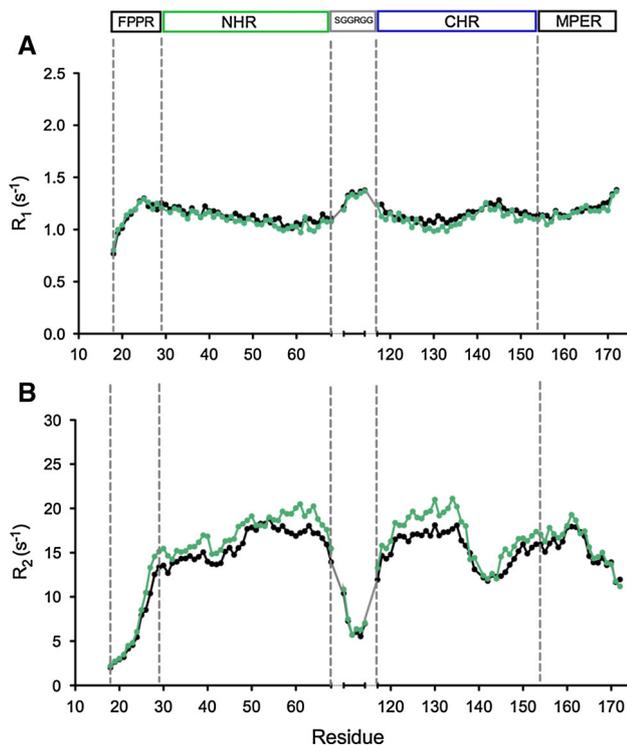


Fig. 6 Comparison of **a** R_1 and **b** R_2 ^{15}N relaxation rates of Core^S₁₇₋₁₇₂ (600 MHz ^1H frequency), recorded at protein (0.4 mM) and DPC (100 mM) concentrations used in Fig. 5 (black symbols), or protein and DPC concentrations of 0.34 and 160 mM, respectively (green symbols). Upon protein dilution and increased detergent concentration, the NHR and CHR ^{15}N R_2 values, on average, increase by 12 %, whereas the R_1 values decrease by 2 %

pH 4.0 and pH 6.0 (Roche et al. 2014). These results highlight the potential active role of the NHR and CHR regions in the fusion mechanism and suggest that the heptad repeat regions may destabilize membranes by directly binding to the lipid bilayer surfaces, thereby contributing to a lowering of the energy barrier associated with membrane fusion.

In such a scenario, the formation of the final 6HB state depends on a competition between intermolecular association of the NHR and CHR helices, including their FPPR and MPER extensions, and membrane binding of these lipophilic regions. Our present study takes advantage of the very high sensitivity of the CSP to track the potential formation of inter-helical interactions in the pre-hairpin state of the fusion process. The very close correspondence of the backbone chemical shifts measured for residues 17–29 of Core^S₁₇₋₁₄₄ and for residues 154–172 of Core^S₃₅₋₁₇₂, relatively to Core^S₁₇₋₁₇₂, clearly points to the absence of significant stabilizing interactions between the FPPR and MPER regions on the surface of DPC micelles (Fig. 3). The absence of interaction between the two membrane proximal regions is confirmed by the ^{15}N relaxation experiments showing nearly identical

heteronuclear $^{15}\text{N}\{-^1\text{H}\}$ NOE values for the various recombinant proteins (Fig. 5), and essentially unchanged internal dynamics in terms of intermediate time scale motions (τ_s , S_s^2) of the helical elements (Table 1).

The absence of intermolecular and long-range intramolecular interactions observed under our experimental conditions, intended to mimic the pre-hairpin states, suggests that interaction between the membrane proximal regions is unlikely to constitute the driving force responsible for the formation of the post-fusion hairpin. In this respect our results differ from those reached in a very recent study, which postulates the formation of a stable NHR-CHR hairpin as an important intermediate (Banerjee and Weliky 2014). In this study, Banerjee and Weliky found that in the absence of detergent, an ectodomain construct corresponding to Core^S₂₄₋₁₅₅ in our nomenclature, is mainly monomeric at pH 3.2 but hexameric at pH 7.4 (Banerjee and Weliky 2014). In contrast to these recent observations, the SEC-MALS experiments reported in our previous study showed that in the absence of detergent, Core^S₃₅₋₁₄₄ is trimeric at both pH 4.0 and 6.0, as expected for a 6HB complex (Roche et al. 2014). In the present study, the SV experiments performed in the absence of detergent at pH 4.0, confirmed the trimeric nature of Core^S₃₅₋₁₄₄, Core^{IL}₃₅₋₁₄₄ and Core^S₁₇₋₁₇₂, but also shows a minor tetramer population for Core^S₁₇₋₁₇₂ (Fig. S2 and Table S1). In our opinion, it therefore is more likely that formation of the final trimeric state may be initiated by interactions between the polar segments of the NHR (S35-N43) and CHR (E136-K144) regions, which lack high membrane affinity but make tight and specific inter-helical contacts in the 6HB (Buzon et al. 2010). Specific interactions between the FP and TM, only accessible after formation of the fusogenic pre-bundle (Fig. 1d), may further stabilize formation of the post-fusion state. Recent experiments performed with FP and TM peptides suggests that interactions between FP and the TM domain of gp41 are functionally important for viral fusion (Reuven et al. 2012).

In the structures of both the pre-fusion gp120-gp41 complex (Julien et al. 2013; Lyumkis et al. 2013) and post-fusion 6HB states (Chan et al. 1997; Weissenhorn et al. 1997; Buzon et al. 2010), the NHR helices form a similar central trimeric helical bundle. Although most current models suggest that this trimeric helical state is maintained during the fusion process (Harrison 2008; Melikyan 2014), early EPR data for the analogous fusion process of the influenza virus involving the hemagglutinin protein suggested that the NHR trimeric bundle dissociates upon interaction with the host cell membrane (Yu et al. 1994), and a similar conclusion based on size exclusion chromatography was reached in a very recent study (Banerjee and Weliky 2014). However, recent reports by Lakomek et al. (2013, 2014) indicate that both gp41¹⁻¹⁹⁴ and gp41²⁷⁻¹⁹⁴

remain trimeric in the presence of DPC, as judged by both analytical centrifugation and SAXS measurement, and confirmed by the long global tumbling time deduced from ^{15}N relaxation studies. By contrast, ^{15}N relaxation data for our constructs that lack the TM domain point to an approximately three-fold shorter global tumbling time, consistent with a monomer. NHR chemical shifts in DPC for all our monomeric constructs are very similar to one another, as well as to those reported for gp41^{1–194} and gp41^{27–194} in DPC. Based on virtually unchanged chemical shifts, the clear absence of interhelical interactions in the previously studied Core^S^{35–144} construct (Roche et al. 2014) therefore excludes the presence of a significant population of a trimeric helical bundle state for the NHR helices of any of the Core^S constructs reported in this study, or in gp41^{1–194} and gp41^{27–194}. Instead, by elimination we conclude that the trimeric state of gp41 must be maintained by its TM segment. Indeed, all constructs that include the TM region show much longer rotational correlation times than proteins lacking this element. Moreover, the TM-containing constructs show the absence of detectable signals for much of the MPER, TM, and the C-terminal half of the CHR regions, indicative of conformational exchange on an intermediate time scale, presumably related to transient intermolecular interactions (MPER and CHR) as well as dynamic behavior of the TM trimeric bundle. Support for the possible presence of intermolecular MPER–MPER interactions can be found in the recent study of Reardon et al. (2014), where a trimeric assembly in the presence of DPC was achieved by linking these MPER elements to the trimeric fusion domain of T4 fibritin.

The presence of a highly conserved GXXXG motif in the TM sequence (G691 and G695), typical of inter-helical interactions in membrane proteins (MacKenzie et al. 1997; Russ and Engelman 2000; Dong et al. 2012), has been interpreted as a signature of the trimeric nature of the gp41 transmembrane region (Kim et al. 2009). Mutagenesis studies have shown that the GXXXG motif, together with the conserved R697 residue, is required for efficient membrane fusion (Kondo et al. 2010; Miyauchi et al. 2010). Oligomerisation of the transmembrane domain has also been observed in the case of the Influenza virus hemagglutinin envelope protein (Tatulian and Tamm 2000) and the Paramyxovirus fusion protein (Smith et al. 2013). The formation of inter-helical interactions within the transmembrane regions might therefore be a general viral strategy to prevent the complete dissociation of the envelope proteins during the course of the fusion. The fact that the resonances of the TM region in gp41^{1–194} gp41^{27–194} are broadened beyond detection Lakomek et al. (2013, 2014) may be caused by the effect of DPC detergent not being an optimal membrane mimetic, or it may be an

intrinsic property of the gp41 trimeric TM arrangement which would bury its three Arg residues deep in the hydrophobic bilayer, and whose dynamic property was found to be important for the kinetics of membrane fusion (Kondo et al. 2010).

Our present study sheds new light on the delicate balance of interactions governing the competition between intermolecular and membrane association during the pre-hairpin stages of the viral fusion process. Our results suggest that inter-helical interactions between the membrane proximal regions do not form the primary driving force for initiating the assembly of the final post-fusion 6HB state of gp41. These regions therefore remain accessible until the very late steps of the fusion pore formation and may represent potential targets for fusion inhibitor drugs. The protein constructs described in our study in the presence of detergent or membranes constitute useful pre-hairpin models to understand the conformational dynamics of gp41, and may prove valuable in the design of future fusion inhibitors.

Acknowledgments We thank Drs. James Baber and Jinfa Ying for technical support and acknowledge support from the Advanced Mass Spectrometry Core of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). This work was funded by the NIH Intramural Research Program of the NIDDK and by the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH.

References

- Banerjee K, Weliky DP (2014) Folded monomers and hexamers of the ectodomain of the HIV gp41 membrane fusion protein: potential roles in fusion and synergy between the fusion peptide, hairpin, and membrane-proximal external region. *Biochemistry*. doi:10.1021/bi501159w
- Bartesaghi A, Merk A, Borgnia MJ, Milne JLS, Subramaniam S (2013) Prefusion structure of trimeric HIV-1 envelope glycoprotein determined by cryo-electron microscopy. *Nat Struct Mol Biol* 20:1352–1357
- Blumenthal R, Durell S, Viard M (2012) HIV entry and envelope glycoprotein-mediated fusion. *J Biol Chem* 287:40841–40849
- Buzon V, Natrajan G, Schibli D, Campelo F, Kozlov MM, Weissenhorn W (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog* 6:7
- Caffrey M, Cai M, Kaufman J, Stahl SJ, Wingfield PT, Covell DG, Gronenborn AM, Clore GM (1998) Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J* 17:4572–4584
- Carr CM, Kim PS (1994) Flu virus invasion: halfway there. *Science* 266:234–236
- Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton N (2007) *Protein NMR spectroscopy: principles and practice*. Elsevier Academic Press, Burlington
- Chan DC, Fass D, Berger JM, Kim PS (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89:263–273
- Chen J, Wharton SA, Weissenhorn W, Calder LJ, Hughson FM, Skehel JJ, Wiley DC (1995) A soluble domain of the membrane-

- anchoring chain of influenza virus hemagglutinin (HA(2)) folds in *Escherichia coli* into the low-pH-induced conformation. *Proc Natl Acad Sci USA* 92:12205–12209
- Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J Am Chem Soc* 112:4989–4991
- Dong H, Sharma M, Zhou H-X, Cross TA (2012) Glycines: role in alpha-helical membrane protein structures and a potential indicator of native conformation. *Biochemistry* 51:4779–4789
- Durrer P, Galli C, Hoenke S, Corti C, Gluck R, Vorherr T, Brunner J (1996) H⁺-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. *J Biol Chem* 271:13417–13421
- Eband RF, Macosko JC, Russell CJ, Shin YK, Eband RM (1999) The ectodomain of HA2 of influenza virus promotes rapid pH dependent membrane fusion. *J Mol Biol* 286:489–503
- Furuta RA, Wild CT, Weng YK, Weiss CD (1998) Capture of an early fusion-active conformation of HIV-1 gp41. *Nat Struct Biol* 5:276–279
- Gallo SA, Finnegan CM, Viard M, Raviv Y, Dimitrov A, Rawat SS, Puri A, Durell S, Blumenthal R (2003) The HIV Env-mediated fusion reaction. *Biochim Biophys Acta* 1614:36–50
- Gao G, Wiczorek L, Peachman KK, Polonis VR, Alving CR, Rao M, Rao VB (2013) Designing a soluble near full-length HIV-1 gp41 Trimer. *J Biol Chem* 288:234–246
- Harrison SC (2008) Viral membrane fusion. *Nat Struct Mol Biol* 15:690–698
- Henderson R (2013) Avoiding the pitfalls of single particle cryo-electron microscopy: einstein from noise. *Proc Natl Acad Sci USA* 110:18037–18041
- Hildinger M, Dittmar MT, Schult-Dietrich P, Fehse B, Schnierle BS, Thaler S, Stiegler G, Welker R, von Laer D (2001) Membrane-anchored peptide inhibits human immunodeficiency virus entry. *J Virol* 75:3038–3042
- Hollmann A, Matos PM, Augusto MT, Castanho MARB, Santos NC (2013) Conjugation of cholesterol to HIV-1 fusion inhibitor C34 increases peptide-membrane interactions potentiating its action. *PLoS ONE* 8:e60302
- Jaroniec CP, Kaufman JD, Stahl SJ, Viard M, Blumenthal R, Wingfield PT, Bax A (2005) Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry* 44:16167–16180
- Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, Klasse PJ, Burton DR, Sanders RW, Moore JP, Ward AB, Wilson IA (2013) Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* 342:1477–1483
- Kim JH, Hartley TL, Curran AR, Engelman DM (2009) Molecular dynamics studies of the transmembrane domain of gp41 from HIV-1. *Biochim Biophys Acta* 1788:1804–1812
- Kliger Y, Peisajovich SG, Blumenthal R, Shai Y (2000) Membrane-induced conformational change during the activation of HIV-1 gp41. *J Mol Biol* 301:905–914
- Kondo N, Miyauchi K, Meng F, Iwamoto A, Matsuda Z (2010) Conformational changes of the HIV-1 envelope protein during membrane fusion are inhibited by the replacement of its membrane-spanning domain. *J Biol Chem* 285:14681–14688
- Korazim O, Sackett K, Shai Y (2006) Functional and structural characterization of HIV-1 gp41 ectodomain regions in phospholipid membranes suggests that the fusion-active conformation is extended. *J Mol Biol* 364:1103–1117
- Lakomek NA, Ying JF, Bax A (2012) Measurement of 15 N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* 53:209–221
- Lakomek N-A, Kaufman JD, Stahl SJ, Louis JM, Grishaev A, Wingfield PT, Bax A (2013) Internal dynamics of the homotrimeric HIV-1 viral coat protein gp41 on multiple time scales. *Angew Chem Int Ed* 52:3911–3915
- Lakomek N-A, Kaufman JD, Stah SJ, Wingfield PT (2014) HIV-1 envelope protein gp41: an NMR study of dodecyl phosphocholine embedded gp41 reveals a dynamic prefusion intermediate conformation. *Structure* 22:1311–1321
- Lev N, Fridmann-Sirkis Y, Blank L, Bitler A, Eband RF, Eband RM, Shai Y (2009) Conformational stability and membrane interaction of the full-length ectodomain of HIV-1 gp41: implication for mode of action. *Biochemistry* 48:3166–3175
- Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* 104:4546–4559
- Lorieau JL, Louis JM, Bax A (2011) Whole-body rocking motion of a fusion peptide in lipid bilayers from size-dispersed 15 N NMR relaxation. *J Am Chem Soc* 133:14184–14187
- Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, Burton DR, Sanders RW, Moore JP, Carragher B, Wilson IA, Ward AB (2013) Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* 342:1484–1490
- MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276:131–133
- Mao Y, Castillo-Menendez LR, Sodroski JG (2013a) Reply to subramaniam, van Heel, and Henderson: validity of the cryo-electron microscopy structures of the HIV-1 envelope glycoprotein complex. *Proc Natl Acad Sci USA* 110:E4178–E4182
- Mao Y, Wang L, Gu C, Herschhorn A, Desormeaux A, Finzi A, Xiang S-H, Sodroski JG (2013b) Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. *Proc Natl Acad Sci USA* 110:12438–12443
- Markosyan RM, Cohen FS, Melikyan GB (2003) HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation. *Mol Biol Cell* 14:926–938
- Melikyan GB (2014) HIV entry: a game of hide-and-fuse? *Curr Opin Virol* 4:1–7
- Melikyan GB, Egelhofer M, von Laer D (2006) Membrane-anchored inhibitory peptides capture human immunodeficiency virus type 1 gp41 conformations that engage the target membrane prior to fusion. *J Virol* 80:3249–3258
- Merk A, Subramaniam S (2013) HIV-1 envelope glycoprotein structure. *Curr Opin Struct Biol* 23:268–276
- Miyauchi, K, Curran, AR, Long, Y, Kondo, N, Iwamoto, A, Engelman, DM, Matsuda, Z (2010) The membrane-spanning domain of gp41 plays a critical role in intracellular trafficking of the HIV envelope protein. *Retrovirology* 7:95
- Ratnayake PU, Sackett K, Nethercott MJ, Weliky DP (2015) pH-dependent vesicle fusion induced by the ectodomain of the human immunodeficiency virus membrane fusion protein gp41: two kinetically distinct processes and fully-membrane-associated gp41 with predominant β sheet fusion peptide conformation. *Biochim Biophys Acta* 1848:289–298
- Reardon PN, Sage H, Dennison SM, Martin JW, Donald BR, Alam SM, Haynes BF, Spicer LD (2014) Structure of an HIV-1-neutralizing antibody target, the lipid-bound gp41 envelope membrane proximal region trimer. *Proc Natl Acad Sci USA* 111:1391–1396
- Reuven EM, Dadon Y, Viard M, Manukovsky N, Blumenthal R, Shai Y (2012) HIV-1 gp41 transmembrane domain interacts with the fusion peptide: implication in lipid mixing and inhibition of virus-cell fusion. *Biochemistry* 51:2867–2878
- Roche J, Louis JM, Grishaev A, Ying J, Bax A (2014) Dissociation of the trimeric gp41 ectodomain at the lipid-water interface suggests an active role in HIV-1 Env-mediated membrane fusion. *Proc Natl Acad Sci USA* 111:3425–3430
- Roux KH, Taylor KA (2007) AIDS virus envelope spike structure. *Curr Opin Struct Biol* 17:244–252

- Russ WP, Engelman DM (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol* 296:911–919
- Sackett K, Shai Y (2002) The HIV-1 gp41 N-terminal heptad repeat plays an essential role in membrane fusion. *Biochemistry* 41:4678–4685
- Sackett K, Nethercott MJ, Shai Y, Weliky DP (2009) Hairpin folding of HIV gp41 abrogates lipid mixing function at physiologic pH and inhibits lipid mixing by exposed gp41 constructs. *Biochemistry* 48:2714–2722
- Sackett K, TerBush A, Weliky DP (2011) HIV gp41 six-helix bundle constructs induce rapid vesicle fusion at pH 3.5 and little fusion at pH 7.0: understanding pH dependence of protein aggregation, membrane binding, and electrostatics, and implications for HIV-host cell fusion. *Eur Biophys J Biophys Lett* 40:489–502
- Sackett K, Nethercott MJ, Zheng Z, Weliky DP (2014) Solid-state NMR spectroscopy of the HIV gp41 membrane fusion protein supports intermolecular antiparallel 13 sheet fusion peptide structure in the final six-helix bundle state. *J Mol Biol* 426:1077–1094
- Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* 69:531–569
- Smith EC, Smith SE, Carter JR, Webb SR, Gibson KM, Hellman LM, Fried MG, Dutch RE (2013) Trimeric transmembrane domain interactions in paramyxovirus fusion proteins. Roles in protein folding, stability, and function. *J Biol Chem* 288:35726–35735
- Subramaniam S (2013) Structure of trimeric HIV-1 envelope glycoproteins. *Proc Natl Acad Sci USA* 110:E4172–E4174
- Tamm LK, Lee J, Liang B (2014) Capturing glimpses of an elusive HIV Gp41 prehairpin fusion intermediate. *Structure* 22:1225–1226
- Tan KM, Liu JH, Wang JH, Shen S, Lu M (1997) Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci USA* 94:12303–12308
- Tatulian SA, Tamm LK (2000) Secondary structure, orientation, oligomerization, and lipid interactions of the transmembrane domain of influenza hemagglutinin. *Biochemistry* 39:496–507
- van Heel M (2013) Finding trimeric HIV-1 envelope glycoproteins in random noise. *Proc Natl Acad Sci USA* 110:E4175–E4177
- Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC (1997) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387:426–430
- Yu YG, King DS, Shin YK (1994) Insertion of a coiled-coil peptide from influenza virus hemagglutinin into membranes. *Science* 266:274–276