Dissociation of the trimeric gp41 ectodomain at the lipid–water interface suggests an active role in HIV-1 Env-mediated membrane fusion

Julien Roche, John M. Louis, Alexander Grishaev, Jinfa Ying, and Adriaan Bax[1]

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Contributed by Adriaan Bax, January 27, 2014 (sent for review December 10, 2013)

The envelope glycoprotein gp41 mediates the process of membrane fusion that enables entry of the HIV-1 virus into the host cell. The actual fusion process involves a switch from a homotrimeric prehairpin intermediate conformation, consisting of parallel coiled-coil helices, to a postfusion state where the ectodomains are arranged as a trimer of helical hairpins, adopting a six-helix bundle (6HB) state. Here, we show by solution NMR spectroscopy that a water-soluble 6HB gp41 ectodomain binds to zwitterionic detergents that contain phosphocholine or phosphatidylcholine head groups and phospholipid vesicles that mimic T-cell membrane composition. Binding results in the dissociation of the 6HB and the formation of a monomeric state, where its two α-helices, N-terminal heptad repeat (NHR) and C-terminal heptad repeat (CHR), become embedded in the lipid–water interface of the virus and host cell. The atomic structure of the gp41 ectodomain monomer, based on NOE distance restraints and residual dipolar couplings, shows that the NHR and CHR helices remain mostly intact, but they completely lose interhelical contacts. The high affinity of the ectodomain helices for phospholipid surfaces suggests that unzippering of the prehairpin intermediate leads to a state where the NHR and CHR helices become embedded in the host cell and viral membranes, respectively, thereby providing a physical force for bringing these membranes into close juxtaposition before actual fusion.

hemagglutinin | HIV-1 fusion inhibitor | RDC | ^{15}N relaxation | chemical shift

The first step of HIV infection involves fusion of the viral and target cell membranes, a process mediated by the viral envelope glycoprotein Env, consisting of subunits gp120 and gp41 (1). The envelope proteins form a noncovalent complex on the viral surface with the trimerized gp41 transmembrane subunit sequestered by three gp120 surface subunits (2–5). 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 result in an extended gp41 conformation (1, 6). In this extended prefusion state, the highly hydrophobic N-terminal fusion peptide (FP) of gp41 anchors in the host cell membrane, while being spatially remote from its transmembrane domain (TM), which traverses the viral membrane (7, 8). After the host cell and viral membranes have fused, the gp41 ectodomain, which links the FP and TM domains, has transitioned into a C3-symmetric six-helix bundle (6HB), with the FP in physical proximity to the TM domain (9). The refolding of gp41 trimers into the highly stable 6HB arrangement is believed to overcome the large free-energy barrier of membrane fusion. Several atomic resolution structures of the 6HB postfusion state have been solved by X-ray crystallography, confirming that the C-terminal heptad repeat (CHR) helices pack in an antiparallel manner into the conserved hydrophobic grooves formed at the surface of the central trimer of N-terminal heptad repeat (NHR) helices (10–12).

Contrary to the postfusion state, structural features of the prehairpin intermediates of HIV-1 gp41 remain the subject of much debate. The functional requirement that gp41's fusion peptide engages the membrane of spatially distant host cells dictates an extended conformation for the time point where FP engages the membrane of the host cell. Cartoon models commonly depict this prehairpin intermediate as an extended trimer of linear NHR and CHR helices (13–17). Recent cryo-EM studies provide more detailed insights into the relatively subtle rearrangement of the trimeric helical NHR core, which is associated with rearrangements of gp120 relative to gp41 on receptor activation of Env, that leads to the release of FP from its hydrophobic burial site at the gp41–gp120 interface (5, 18, 19). Subsequent dissociation of the gp120 subunits leaves the gp41 core in a state somewhat similar to the common cartoon models, lacking the trimer-stabilizing interactions supplied by gp120.

Although it seems clear that, initially, gp41 directly engages the viral and host cell membranes only by means of its TM and FP domains, there is evidence that, subsequently, the NHR region also interacts directly with the membranes and actively participates in the fusion process. In particular, the NHR-derived peptide, N36, binds to both zwitterionic and negatively charged phospholipid vesicles (20), whereas the N70 peptide, which encompasses the FP and NHR domains, is four times more fusogenic than FP alone for negatively charged membranes (21). The latter result suggests that the NHR segment takes an active role in destabilizing membranes and works synergistically with FP to increase the efficiency of lipid mixing. In another elegant set of experiments, Wexler-Cohen and Shai (14) showed that NHR-mimicking peptides, designed to interfere with formation of gp41's 6HB state by competing with gp41 NHR insertion into

Significance

Infection by HIV-1 requires fusion of viral and host cell membranes, a process mediated by viral protein gp41. Although extensive structural detail on both pre- and postfusion gp41 states is available from X-ray crystallography and cryo-EM studies, little is known about the actual transition. This NMR study of a trimeric gp41 ectodomain, which connects viral and host cell membranes in the prefusion state, suggests a fusion model, where this domain unzippers from opposite ends because of the affinity of its two α-helices for viral and host cell membranes. In this model, the change in orientation of the ectodomain helices, which is associated with membrane binding, provides the driving force that pulls the membranes into the close juxtaposition required for fusion.

Author contributions: J.R., J.M.L., and A.B. designed research; J.R., J.M.L., and J.Y. performed research; J.M.L. contributed new reagents/analytic tools; J.R., J.M.L., and A.G. analyzed data; and J.R. and A.B. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The NMR, atomic coordinates, chemical shifts, and restraints have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2MK3).

[1]To whom correspondence should be addressed. E-mail: bax@nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1401397111/-/DCSupplemental.

the 6HB, have strongly increased inhibitory activity when they carry a membrane-anchoring alkyl chain. Increased inhibition is seen regardless of whether the alkyl chain is attached at the N or C terminus of the NHR peptide, suggesting that the gp41 NHR domain is embedded in the membrane surface. 6HB oligomers formed by NHR- and CHR-derived synthetic peptides dissociate in the presence of either zwitterionic or negatively charged phospholipid vesicles (20, 22). 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 (23).

In the present study, we show that the 6HB complex formed by an ectodomain that contains large segments of the NHR and CHR helices, connected by a six-residue linker (Core^{S}), dissociates and forms stable monomers on binding to either dodecyl phosphocholine (DPC) micelles or phospholipid vesicles of a lipid composition that mimics the T-cell membrane. The transition from trimers to monomers is associated with a significant decrease in α-helicity and also observed for a longer ectodomain construct (Core^{IL}) that encompasses the native immunodominant loop (IL) connecting the NHR and CHR helices. The Core^{S} construct was chosen for detailed characterization of the structure and dynamics of the gp41 ectodomain monomer in the presence of DPC micelles. An atomic structure determination by NMR spectroscopy of the gp41 ectodomain monomer, based on residual dipolar coupling (RDC) and NOE restraints, reveals a monomeric, flexibly linked two-helical structure lying on the surface of the DPC micelle without any specific interaction between the stable and well-defined NHR and CHR helices. We propose that formation of this lipid-bound state, where CHR embeds in the viral membrane and NHR in the membrane of the host cell, provides the force for pulling the two membranes into close juxtaposition, thereby priming the system for membrane fusion. After fusion, close spatial proximity between the opposite ends of the ectodomain then permits their tight interaction, which is seen in 6HB crystal structures of the full-length gp41 ectodomain (9).

Results

Secondary Structure and Oligomeric State of gp41 Ectodomain. We expressed and purified a recombinant protein, Core^{S}, containing the NHR and CHR segments connected by a 6-residue linker (L6) (Fig. 1*A*), which is known to form a stable 6HB homotrimeric complex in aqueous solution (11). CD spectra of Core^{S} recorded at pH 4.0 show the characteristic signature of an α-helical protein with a deep minimum at 222 nm (Fig. 1*B*), corresponding to *ca.* 83% helical content. The addition of 10 mM DPC results in a 23% loss in helicity (Fig. 1*B*), indicating a substantial structural perturbation of Core^{S} on binding to the DPC micelle. The same change of the CD spectrum is observed when Core^{S} is mixed with dihexanoyl phosphatidylcholine (DHPC) micelles (Fig. S1*A*), which contrasts with virtually no change of the CD spectrum on addition of the detergents 3-([3-cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) (Fig. S1*B*) or lauryl maltose neopentyl glycol (MNG-3) (Fig. S1*C*), suggesting that the presence of phospholipid head groups is important for the binding of Core^{S}.

CD spectra of Core^{S} were also recorded at pH 6.0, showing the same decrease in helicity on addition of DPC at pH 4.0 (Fig. S1*E*). Importantly, a very similar perturbation is observed when Core^{S} is mixed with vesicles known as LM3, which mimic the T-cell membrane lipid composition (24) (Fig. S1*D*).

Using size-exclusion chromatography coupled to multiangle light scattering, refractive index, and UV measurements (SEC-MALS), we find that the secondary structure perturbation described above is correlated with a change in the oligomeric state of Core^{S}. For a 5 μM protein solution in the absence of detergent, a single elution peak corresponding to the Core^{S} trimer (molecular mass = 30.6 kDa) is observed, which was expected for a stable 6HB trimer (Fig. 1*C*, blue trace). By contrast, in the presence of 10 mM DPC, the SEC-MALS data show an elution peak corresponding to monomeric Core^{S} (molecular mass = 11 kDa) bound to a DPC micelle (molecular mass = 18.5 kDa) (Fig. 1*C*, red trace). No elution peak corresponding to the trimeric form is observed under such conditions, indicating that an excess of DPC micelles completely shifts the equilibrium to the monomeric state of Core^{S}. SEC-MALS measurements were also performed for Core^{S} at pH 6.0, again showing a complete trimer-to-monomer transition in the presence of DPC and molecular masses for the trimer and the micelle-bound monomer very similar to the masses seen at pH 4.0 (Fig. S1*F*).

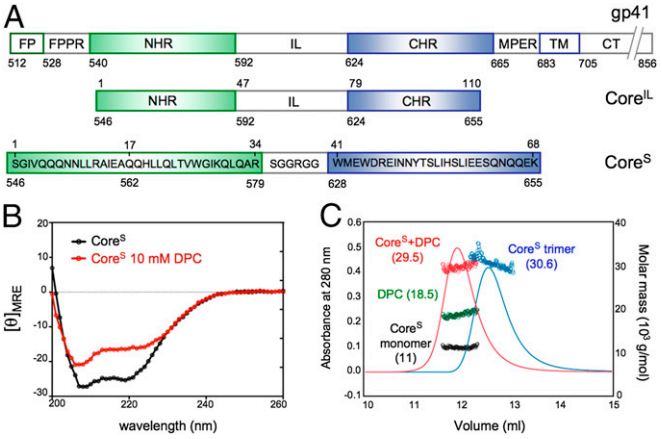

Fig. 1. Sequence and properties of gp41. (*A*) Schematic representation of the gp41 sequence, including the FP, FPPR, NHR, IL, CHR, MPER, TM, and intraviral C-terminal domain (CT). The constructs used in the present study contain the NHR and CHR segments connected by either IL or L6. The numbering 512–704 refers to the Env precursor sequence, whereas the 1–68 numbering is used for Core^{S}. In addition, the Core^{S} sequence contains four extra residues (GSHM) at its N terminus, which correspond to an uncleaved fragment of the original tag (*SI Materials and Methods*). (*B*) CD spectra of Core^{S}, reported here as the mean residue ellipticity (10^3 degrees centimeter^2 decimoles^{-1} residue^{-1}), were recorded in the absence of detergent (black) and the presence of 10 mM DPC (red) at 310 K. (*C*) Molecular mass analysis of Core^{S} (including its N-terminal His-tag) (*SI Materials and Methods*) in the absence and presence of DPC as determined by SEC-MALS. The elution profiles monitored by the absorbance at 280 nm are shown for the Core^{S} trimer (blue, 30.6 ± 0.3 kDa) and a Core^{S} monomer bound to a DPC micelle (red, 29.5 ± 0.4 kDa), with the DPC micelle contribution in green (18.5 ± 0.4 kDa) and the Core^{S} monomer in black (11 ± 0.1 kDa).

To exclude that the trimer-to-monomer transition observed for Core^{S} in the presence of DPC is a consequence of substituting the native immunodominant loop (IL) by L6, the same measurements were repeated for a construct that included IL instead of L6 (Fig. 1*A*). CD measurements on Core^{IL} show a similar decrease in helical content on addition of 10 mM DPC (Fig. S1 *G* and *I*), whereas the SEC-MALS data again indicate that the trimeric population of Core^{IL} undergoes a complete shift to a monomeric micelle-bound state in the presence of DPC (Fig. S1 *H* and *J*). These results, therefore, confirm that the trimer-to-monomer transition is a common property of both Core^{IL} and Core^{S} and not a simple consequence of the replacement of the IL region by a short linker.

Structure and Dynamics of the Trimeric and Monomeric States of Core^{S}. The structure and backbone dynamics of the trimeric and monomeric forms of Core^{S} were studied by solution NMR spectroscopy. In the absence of DPC, the ^{1}H-^{15}N TROSY-HSQC spectrum of Core^{S} at pH 4.0 presents all of the characteristics of a stably folded protein, with 68 well-dispersed amide chemical shifts and uniform resonance line widths, indicating that the trimer

is C3-symmetric. The ^{1}H-^{15}N TROSY-HSQC spectrum recorded for the Core^{S} monomer in the presence of 100 mM DPC (Fig. 2*A*) also shows 68 well-dispersed amide resonances but with large chemical shift differences relative to the trimer (Table S1). ^{1}H-^{15}N TROSY-HSQC spectra, recorded at pH 6.0 in the absence and presence of 100 mM DPC, show the same characteristic chemical shift differences of the trimer-to-monomer transition (Fig. S2 *A* and *B*). ^{1}H-^{15}N TROSY-HSQC spectra were also recorded in the presence of LM3 vesicles at both pH 4.0 and 6.0 (Fig. S2 *C* and *D*), showing the disappearance of Core^{S} cross-peaks for all ordered residues in the slowly tumbling lipid-bound state; this observation confirmed that Core^{S} binds to these T cell-mimicking vesicles, while exhibiting resonances for the observable, dynamically disordered residues that fall close to the positions seen in the monomeric DPC-solubilized construct.

The secondary chemical shifts of the $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and $^{13}C'$ nuclei, which refer to the difference between the observed chemical shifts and the corresponding residue-specific random coil values, are sensitive indicators of local secondary structure. With the exception of the two N-terminal and the six C-terminal residues, large positive $^{13}C^{\alpha}$ secondary chemical shifts are observed for the NHR and CHR regions in the absence of detergent, which is indicative of α-helical structure. The six residues composing the artificial linker between the NHR and CHR helices show chemical shifts close to random coil chemical shift values, indicating that this linker is dynamically disordered in solution (Fig. 2*B*). The $^{13}C^{\alpha}$ secondary chemical shifts measured for the Core^{S} monomer in the presence of 100 mM DPC also show large positive values for most of the CHR region, but significant differences relative to the shifts of the Core^{S} trimer are seen in the NHR segment, particularly for residues I3–N9, which exhibit decreased deviations from random coil values in the monomeric state (Fig. 2*B*). In addition, the very small and even negative values observed for A16 and Q17, respectively, point to the presence of a break in the NHR helix of the Core^{S} monomer.

Measurement of the ^{15}N spin–lattice (R_1) and spin–spin (R_2) relaxation rates together with the heteronuclear ^{15}N-{^{1}H} NOE, which was carried out in both the absence and presence of 100 mM DPC, permits quantitative evaluation of the backbone dynamics of Core^{S} in the two states. With the exception of the eight N- and seven C-terminal residues, highly uniform NOE values (between 0.81 and 0.85) that fall close to their theoretical rigid limit are observed in the absence of DPC for both the NHR and CHR helices of Core^{S} (Fig. 2), indicating that they adopt a highly ordered conformation in the trimer. Uniform R_2 and R_1 relaxation rates of *ca.* 15 and 1 s^{-1}, respectively (Fig. S3), correspond to a rotational correlation time of 9.9 ± 0.2 ns, close to the value expected for a 24.9 kDa globular protein. Addition of DPC resulted in a small but rather uniform decrease of the heteronuclear NOE (Fig. 2). Together with decreased R_2 relaxation rates and increased R_1 values (Fig. S3), these data indicate that both the NHR and CHR helices tumble more rapidly in the detergent-attached monomeric state. A model free analysis (26) of the ^{15}N relaxation data additionally reveals increased internal motions for the NH vectors of the first nine N-terminal residues of the Core^{S} monomer, with S^2 order parameters ranging between 0.4 and 0.7 (Fig. S3*C* and Table S2). Comparison of the R_2 relaxation rates measured at 600 and 800 MHz (Fig. S3*D*) shows no evidence of slow (on the microsecond to millisecond timescale) conformational exchange. Therefore, any additional dynamics that may be present on a timescale slower than the overall molecular tumbling, which escapes the Lipari–Szabo relaxation analysis, must take place on a timescale faster than *ca.* 30 μs.

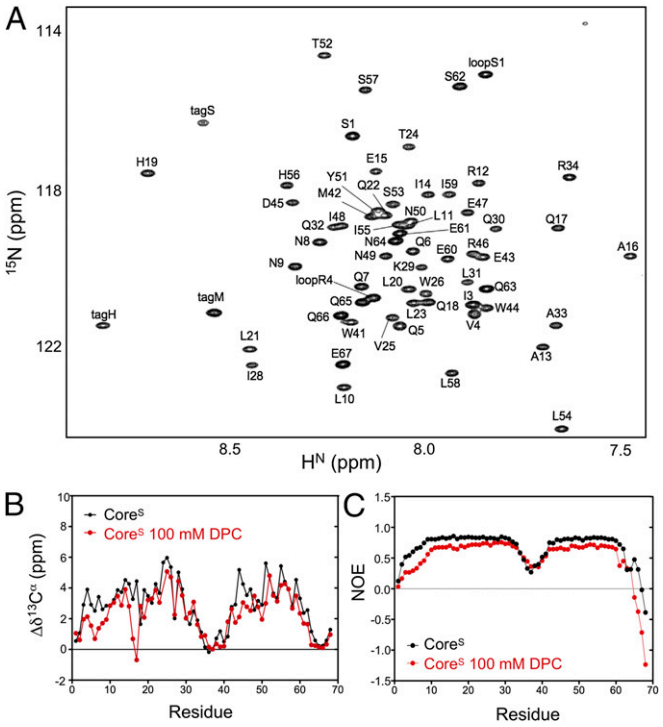

Fig. 2. NMR characterization of Core^{S} at pH 4.0 and 310 K. (*A*) The most crowded region of the ^{1}H-^{15}N TROSY-HSQC spectrum in the presence of 100 mM DPC. Assignments are included in Table S1. (*B*) Comparison of the secondary $^{13}C^{\alpha}$ chemical shifts ($\Delta\delta C^{\alpha}$) of Core^{S} in the absence (black) and presence (red) of 100 mM DPC (red). $\Delta\delta C^{\alpha}$ values represent the difference between the measured $^{13}C^{\alpha}$ chemical shifts and the temperature- and pH-corrected random coil values (25). (*C*) Steady state heteronuclear ^{15}N-{^{1}H} NOE values of Core^{S} (600 MHz ^{1}H frequency) in the absence of detergent (black) and the presence of 100 mM DPC (red).

Comparison of the Core^{S} Trimer with 6HB Crystal Structures. An X-ray structure of Core^{S}, crystallized in the absence of detergent (11), revealed the same 6HB structure seen for the ectodomain of other class I viral fusion proteins. RDCs are exquisitely sensitive probes for evaluating how close this structure resembles the structure present in aqueous solution. Therefore, we collected a nearly complete set of $^{1}D_{NH}$, $^{1}D_{NC'}$, $^{2}D_{HNC'}$, and $^{1}D_{C\alpha C'}$ RDCs and fitted these couplings to the X-ray structure (Protein Data Bank ID code 1SZT) (11), yielding a Q factor of 0.274. Considering the limited crystallographic resolution at which the coordinates had been determined (2.4-Å resolution), this Q factor indicates good agreement with the crystal structure, with no obvious outliers or systematic differences (Fig. S4*A*). A higher-resolution crystal structure of Core^{S}, crystallized in the presence of n-octyl-β-glucopyranoside (Protein Data Bank ID code 1DF4; 1.45-Å resolution), had been interpreted as an alternative, micelle-bound conformation of the 6HB structure; however, the C^{α} coordinate rmsd between the 1SZT and 1DF4 structures is only modest (0.58 Å) (27). Interestingly, a fit of our experimental RDCs measured in the absence of detergent to the 1DF4 structure actually shows significantly better agreement, with a Q factor = 0.192 (Fig. S4*B*), than to the detergent-free structure. This result indicates that the fit of the experimental RDCs to 1SZT is principally limited by the accuracy of the atomic coordinates derived from a 2.4-Å map and that 1DF4 actually represents a more accurate coordinate representation of the detergent-free 6HB conformation. Importantly, these results confirm that Core^{S} in solution adopts the same trimeric 6HB structure seen in the crystalline state.

Structure of Monomeric Core^{S}. The structure of Core^{S} in the micelle-bound monomeric state was derived by simulated annealing

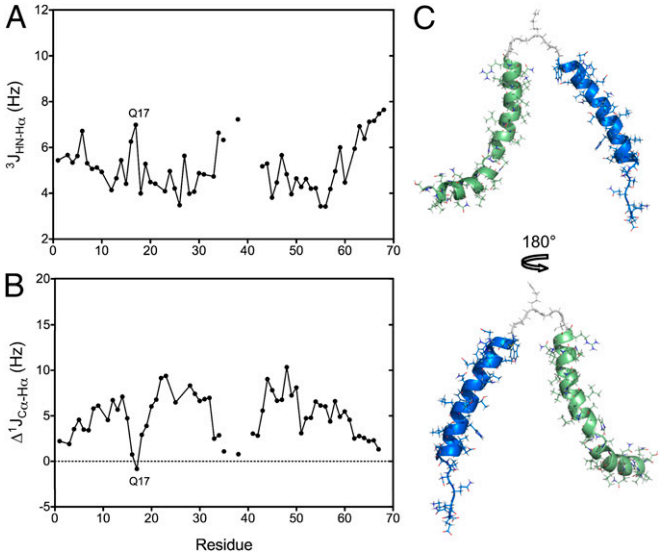

Fig. 3. Core^S in the presence of DPC. (*A*) $^3J_{HN\text{-}H\alpha}$ couplings and (*B*) secondary $\Delta^1J_{C\alpha-H\alpha}$ values reporting on secondary structure. $\Delta^1J_{C\alpha-H\alpha}$ is the difference between the measured $^1J_{C\alpha-H\alpha}$ coupling and the residue-specific random coil value. Values measured for residues within the NHR and CHR segments are connected by solid lines for visual purposes. (*C*) Structure of Core^S in the micelle-bound monomeric state, with the NHR helix in green, the CHR in blue, and the flexible linker (residues 35–40) in gray. Although the fourfold degeneracy in the average relative orientations of NHR and CHR, intrinsic to RDC analysis (29), is broken by the requirement that both helices adhere with their lipophilic surface to the same micelle, their instantaneous relative orientation is subject to dynamic disorder. Without direct interhelical contacts, translationally, the relative position of the two helices also is ill-defined. For structural statistics, see Table S3.

calculations based on 465 sequential and short-range NOE distance restraints, 114 backbone dihedral angle restraints, and 157 backbone RDCs. The prediction of the backbone dihedral angles provided by the TalosN program (28) and based on the experimental ^{15}N, $^1H^N$, $^1H^\alpha$, $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ chemical shifts was complemented by the measurements of $^3J_{HN\text{-}H\alpha}$ and $^1J_{C\alpha\text{-}H\alpha}$ scalar couplings (Fig. 3). The two residues for which very small and even negative secondary $^{13}C^\alpha$ chemical shifts were measured, A16 and Q17 (Fig. 2*B*), also show nonhelical $^3J_{HN\text{-}H\alpha}$ couplings (6.3 and 7.0 Hz, respectively) (Fig. 3*A*) and reduced secondary $^1J_{C\alpha\text{-}H\alpha}$ couplings (0.75 and −0.84 Hz, respectively) (Fig. 3*B*), confirming the nonhelical backbone torsion angles of these two residues. The somewhat extended conformations of A16 and Q17 result in a clear kink of the NHR helix (Fig. 3*C*), with the N-terminal segment at an angle of *ca.* 60° relative to the main axis of the helical segment, which is formed by residues 18–34. The CHR helix is largely preserved in the monomeric state, except for its eight most C-terminal residues. These C-terminal residues all have polar side chains, and their inability to engage the lipid surface in an α-helical conformation, therefore, is (not surprising) resulting in dynamic disorder. Analogously, five sequential polar residues near the N terminus, QQQNN, prevent lipid binding of this section of the NHR, also resulting in dynamic disorder and only transient helical character, which was judged by $^{13}C^\alpha$ secondary shifts and $^3J_{HNH\alpha}$ and $^1J_{C\alpha H\alpha}$ couplings (Fig. 3).

Absence of NHR and CHR Interaction in the Monomeric State. Although no interhelical interactions were observed in the 2D and 3D NOESY spectra of Core^S in the presence of DPC, weak transient interactions cannot be excluded a priori, because the corresponding NOEs would be notoriously difficult to detect. However, chemical shifts are exquisitely sensitive to even transient, weak interactions. To investigate the presence of potential weak interactions between NHR and CHR helices, we compared the chemical shifts observed for the 68-residue Core^S monomer with chemical shifts recorded for separately purified recombinant forms of the NHR and CHR peptides. CD spectra recorded in 50 mM sodium acetate (pH 4.0) at 310 K show that the NHR peptide is intrinsically disordered in the absence of detergent but adopts an α-helical conformation on addition of DPC (Fig. S5*A*). Comparison of the chemical shifts of Core^S in the presence of DPC with the shifts of the two separate peptide samples shows them to be essentially indistinguishable for not only $^{13}C^\alpha$ (Fig. 4*A*) but also, the amides $^1H^N$ (Fig. S5*C*) and ^{15}N (Fig. S5*D*). The minor differences observed for the terminal regions of NHR and CHR reflect the presence of additional residues that extend the isolated NHR and CHR peptides, necessary for their isolation (*SI Materials and Methods*). Other than these minor differences, the very close correspondence between the chemical shifts measured for the isolated peptides and the micelle-associated Core^S indicates that the NHR–CHR interactions are completely disrupted in the monomeric state of Core^S. In addition, the very close correspondence between the $^{13}C^\alpha$ secondary chemical shifts measured for the two peptides and the shifts measured for Core^{IL} (Fig. 4*B*) suggests that the interhelical interactions are also disrupted in this longer construct. The absence of stabilizing interactions between the two helices together with the high flexibility of the interhelical linker suggest that the relative orientation and position on the lipid surface are subject to large dynamic disorder, with the structure depicted in Fig. 3*C* only representing an average view. Indeed, when immersed in anisotropically compressed acrylamide gel, the alignment strength of the larger N-terminal helix is found to be greater than for the shorter C-terminal helix, confirming their dynamic relative arrangement.

Core^S at the Phospholipid–Water Interface. Paramagnetic relaxation enhancement has become a standard method to study the partitioning of peptides and proteins at the water–phospholipid interface. The solvent- and micelle-associated surfaces of Core^S in the monomeric state were identified by comparing the amide signal attenuation induced by two paramagnetic agents: 16-doxylstearic acid (16-DSA), which is confined to the hydrophobic interior of the DPC micelle, and gadodiamide (Omniscan), which remains free in solution. ^1H-^{15}N TROSY-HSQC spectra of Core^S with 100 mM DPC were recorded in the presence of either 2 mM 16-DSA or 2 mM Omniscan, and the attenuation profile of each amide group was calculated by comparing the cross-peak intensities in the presence and absence of paramagnetic agent

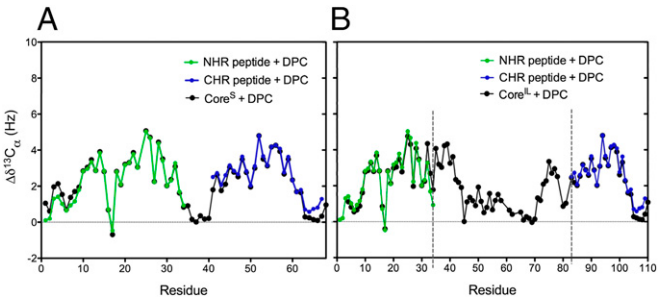

Fig. 4. Comparison of the secondary $^{13}C^\alpha$ chemical shifts of the individual NHR (green) and CHR (blue) peptides with shifts of (*A*) Core^S (black) and (*B*) Core^{IL} [black; all in 50 mM sodium acetate (pH 4.0) in the presence of 100 mM DPC]. The individual NHR and CHR helix constructs contain additional residues at their N and C termini (*SI Materials and Methods*), respectively, that are not present in Core^S and presumed to be responsible for the small chemical shift differences near the termini, which seem dynamically disordered in the two peptides as well as in Core^S and Core^{IL}.

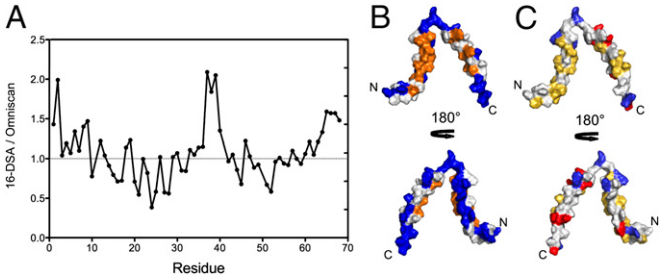

Fig. 5. Probing of the interaction between monomeric Core^S and the phospholipid interface. (*A*) Ratios of attenuation induced by the paramagnetic agents, 16-DSA and Omniscan, as a function of residue number. The attenuation of each amide signal in the ^{1}H-^{15}N TROSY-HSQC spectra of Core^S with 100 mM DPC was independently determined on addition of either 2 mM 16-DSA or 2 mM Omniscan (Table S2). (*B*) Surface representation of the Core^S monomer structure, with the residues showing the largest 16-DSA/Omniscan ratios colored in blue (solvent-exposed) and the smallest ratios colored in orange (micelle-exposed). (*C*) Surface representation of the Core^S monomer colored on the basis of residue type (blue, positively charged; red, negatively charged; yellow, hydrophobic). The N and C termini are marked N and C, respectively.

(Table S2). The ratio between the attenuations induced by 16-DSA and Omniscan (16-DSA/Omniscan) (Fig. 5*A*) reveals which amide groups are more affected by Omniscan than 16-DSA (ratio > 1; therefore, closer to solvent) or micelle-exposed (ratio < 1). The largest and smallest 16-DSA/Omniscan ratios are marked on the structure of the Core^S monomer in blue and orange, respectively, in Fig. 5*B*, and they show a clear partitioning of the solvent- and micelle-exposed surfaces.

Discussion

Before reaching the postfusion 6HB state, the NHR and CHR regions have the opportunity to interact with their adjacent membranes, and a growing body of evidence suggests that these heptad repeat regions may play an active role in destabilizing membranes by directly binding to the lipid bilayers (20–23). Our study shows that the 6HB trimeric structure of a recombinant ectodomain, lacking the membrane-interacting domains FP, FP proximal region (FPPR), membrane proximal external region (MPER), and TM, dissociates into stable monomers on binding to zwitterionic detergent micelles and behaves analogously in the presence of vesicles that mimic the T-cell membrane composition. Although the ability of gp41 constructs lacking the MPER and TM regions to induce lipid mixing and vesicle fusion has been shown to depend strongly on pH (30), we find that dissociation into monomers occurs both at pH 4.0 and 6.0. In the lipid-bound state, both the NHR and CHR helices are embedded at the water–lipid interface, thereby destabilizing their respective membranes and lowering the barrier for membrane fusion (31). The trimer-to-monomer transition introduces a significant kink in the NHR helix at residue Q562 (Q17 in Core^S numbering). The force needed to maintain the lipid-bound NHR in its kinked conformation can only be provided by its interaction with the lipids and therefore, must be accompanied by additional destabilization of the lipid interface, thereby also contributing to a lowering of the energy barrier associated with membrane fusion. In this respect, we note that, although the small degree of helical axis curvature observed for the NHR and CHR Core^S helices in our detergent-solubilized model system is likely to differ from any curvature present when these helices are embedded in the host cell and viral membranes, the kink in the micelle-bound NHR helix seems to be an essential attribute for its binding to a contiguous hydrophobic bilayer surface. This kink is necessary to avoid the polar side chain of residue Q562 from facing the bilayer interior, while allowing the hydrophobic side chains of L555, L556, and I559, which precede the kink, to engage the bilayer simultaneously with the side chains of L565, L566, V570, and I573.

Recently, we found that the CHR, MPER, and TM regions of a much longer gp41 construct (residues 512–705; see Fig. 1*A*) are subject to extensive conformational exchange processes in the presence of DPC (32). Although sedimentation equilibrium centrifugation and SEC-MALS data unambiguously showed that this gp41$^{512–705}$ remains trimeric under such conditions, the chemical shifts of the NHR residues in gp41$^{512–705}$ correlate much closer with the shifts of the Core^S monomer (R^2 = 0.76 and 0.74 for the $^1H^N$ and secondary $^{13}C^\alpha$ chemical shifts, respectively) than with the Core^S trimer [R^2 = 0.24 (δ^1H^N) and R^2 = 0.55 ($\Delta\delta^{13}C_\alpha$)] (Fig. S6). This paradox is resolved by recognizing that the membrane-associated TM region is responsible for retaining the trimeric state of gp41$^{512–705}$, even in the absence of stable NHR–CHR interhelical interactions. The present Core^S data indicate that the affinity of the NHR and CHR segments for phospholipid surfaces is strong enough to break the thermodynamically very stable 6HB trimeric state and therefore, more than sufficient to disrupt the much weaker intermolecular NHR and potential CHR interactions in the initial extended prehairpin intermediate state (Fig. 6*A*), thereby rapidly transitioning to a collapsed state (Fig. 6*B*). This transition pulls the viral and host cell membranes closer to one another to a distance that is limited by the length of the IL (Fig. 1*A*). Considering that a significant segment of IL (residues I580–D589 after NHR and S618–T627 preceding CHR) also is lipophilic and α-helical (Fig. 4), the effective intermembrane distance likely is even shorter and also dependent on the oxidation state of the two Cys residues (C598 and C604) located in the nonlipophilic segment of IL.

Destabilization of the viral and target cell membranes introduced by the heptad helices and the FP in the collapsed prehairpin intermediate state (Fig. 6*B*) coupled with their spatial proximity then creates a state conducive to the formation of a hemifusion stalk, in which the outer leaflets of the viral and host cell membranes are fused (31), and which can progress to formation of a small fusion pore (Fig. 6*C*). Taking advantage of the temperature dependence of the fusion pore growth, Markosyan

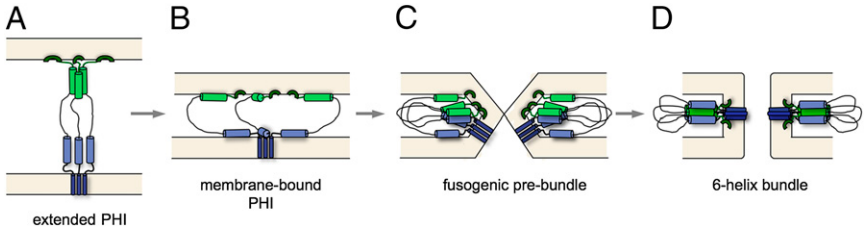

Fig. 6. Model of the intermediate steps in gp41-driven fusion of the viral and target cell membranes showing the NHR (light green) and CHR (light blue) segments and the membrane-anchoring elements [FP (dark green) and TM (dark blue)] at four different stages of the fusion process. (*A*) The short-lived extended prehairpin intermediate (PHI) state, where both TM and NHR are presumed responsible for maintaining the trimeric nature. (*B*) The collapsed PHI state, where NHR and CHR have become embedded in the viral and host cell membranes, thereby pulling the membranes into juxtaposition. (*C*) Formation of fusogenic prebundles, which is possibly initiated by contacts between the short polar segments at opposing ends of the NHR and CHR. (*D*) Formation of mature, postfusion 6HB trimers, which are stabilized by FP–TM, FPPR–MPER, and 6HB NHR–CHR interactions.

et al. (33) have shown that folding of the 6HB is not complete until the very late stage of pore formation. Based on this kinetic argument, we propose that the initial formation of fusion pores is driven by the association of gp41 trimers into prebundle complexes (Fig. 6*C*). Formation of these complexes then 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 (32, 34). Bundling of MPER with FPPR residues stabilizes the 6HB state (9) but only becomes kinetically accessible after these regions have progressed to a state of close spatial proximity (Fig. 6*C*). We speculate that formation of this trimeric state may be initiated by interactions between the polar segments of the NHR (S546–N554) and CHR (E647–K655) regions, which lack high membrane affinity but make tight and specific interhelical contacts in the 6HB. Specific interactions between the FP and TM (35), which are only accessible after formation of the fusogenic prebundle, may further stabilize formation of the postfusion state. Competition between intermolecular and membrane association may also be impacted by a shift in lipid composition after outer leaflet lipids of the viral and host cells can mix with one another by translational diffusion through the hemifusion stalk, which would be a slow and strongly temperature-dependent process.

In our model, the fusogenic prebundle complexes, comprising both membrane- and self-associated heptad regions, represent the actual target for the peptides used for fusion inhibition (15, 16, 36). Such long-lived prebundle conformations would also represent an ideal target for the membrane-conjugated class of inhibitory NHR- and CHR-mimicking peptides (37–39) as well as neutralizing antibodies, which can tightly engage incomplete states of the 6HB core (40, 41).

Materials and Methods

Core^S, identical in sequence to the N34-L6-C28 construct described in ref. 11, was expressed with an N-terminal His-tag to aid its purification. The N-terminal nonnative residues were removed by thrombin cleavage, with the exception of four residues (GSHM) remaining at the N terminus of the Core^S sequence, which was followed by size exclusion chromatography under denaturing conditions and reverse-phase HPLC.

NMR measurements were carried out at 500, 600, 800, and 900 MHz on uniformly $^2H/^{15}N/^{13}C$-, $^{15}N/^{13}C$-, and $^2H/^{15}N$-enriched samples at protein concentrations of *ca.* 0.5 mM (monomer) in both the absence and presence of 100 mM DPC. The NMR structure of the Core^S monomer was calculated using NOE distance restraints, RDCs, and TalosN dihedral restraints (28) using X-PLOR-NIH v2.34 (42). Details are in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Annie Aniana for help with protein expression and purification and Robert Blumenthal and Leonid Chernomordik for helpful discussions, and we acknowledge support from the National Institute of Diabetes and Digestive and Kidney Diseases MS facility. This work was funded by the National Institutes of Health Intramural Research Programs of the National Institute of Diabetes and Digestive and Kidney Diseases and the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, National Institutes of Health.

1. Harrison SC (2008) Viral membrane fusion. *Nat Struct Mol Biol* 15(7):690–698.
2. Roux KH, Taylor KA (2007) AIDS virus envelope spike structure. *Curr Opin Struct Biol* 17(2):244–252.
3. Merk A, Subramaniam S (2013) HIV-1 envelope glycoprotein structure. *Curr Opin Struct Biol* 23(2):268–276.
4. Julien JP, et al. (2013) Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* 342(6165):1477–1483.
5. Lyumkis D, et al. (2013) Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* 342(6165):1484–1490.
6. Furuta RA, Wild CT, Weng YK, Weiss CD (1998) Capture of an early fusion-active conformation of HIV-1 gp41. *Nat Struct Biol* 5(4):276–279.
7. Gallo SA, et al. (2003) The HIV Env-mediated fusion reaction. *Biochim Biophys Acta* 1614(1):36–50.
8. Blumenthal R, Durell S, Viard M (2012) HIV entry and envelope glycoprotein-mediated fusion. *J Biol Chem* 287(49):40841–40849.
9. Buzon V, et al. (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog* 6(5):e1000880.
10. Chan DC, Fass D, Berger JM, Kim PS (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89(2):263–273.
11. 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(23):12303–12308.
12. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC (1997) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387(6631):426–430.
13. Weissenhorn W, et al. (1999) Structural basis for membrane fusion by enveloped viruses. *Mol Membr Biol* 16(1):3–9.
14. Wexler-Cohen Y, Shai Y (2009) Membrane-anchored HIV-1 N-heptad repeat peptides are highly potent cell fusion inhibitors via an altered mode of action. *PLoS Pathog* 5(7):e1000509.
15. Eggink D, Berkhout B, Sanders RW (2010) Inhibition of HIV-1 by fusion inhibitors. *Curr Pharm Des* 16(33):3716–3728.
16. Ashkenazi A, Shai Y (2011) Insights into the mechanism of HIV-1 envelope induced membrane fusion as revealed by its inhibitory peptides. *Eur Biophys J* 40(4):349–357.
17. Garg H, Viard M, Jacobs A, Blumenthal R (2011) Targeting HIV-1 gp41-induced fusion and pathogenesis for anti-viral therapy. *Curr Top Med Chem* 11(24):2947–2958.
18. 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(12):1352–1357.
19. Tran EEH, et al. (2012) Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. *PLoS Pathog* 8(7):e1002797.
20. 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(5):1103–1117.
21. Sackett K, Shai Y (2002) The HIV-1 gp41 N-terminal heptad repeat plays an essential role in membrane fusion. *Biochemistry* 41(14):4678–4685.
22. Kliger Y, Peisajovich SG, Blumenthal R, Shai Y (2000) Membrane-induced conformational change during the activation of HIV-1 gp41. *J Mol Biol* 301(4):905–914.
23. Lev N, et al. (2009) Conformational stability and membrane interaction of the full-length ectodomain of HIV-1 gp41: Implication for mode of action. *Biochemistry* 48(14):3166–3175.
24. Aloia RC, Tian HR, Jensen FC (1993) Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes. *Proc Natl Acad Sci USA* 90(11):5181–5185.
25. Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: Effects of temperature and pH. *J Biomol NMR* 49(2):139–149.
26. 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(17):4546–4559.
27. Shu W, Ji H, Lu M (2000) Interactions between HIV-1 gp41 core and detergents and their implications for membrane fusion. *J Biol Chem* 275(3):1839–1845.
28. Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56(3):227–241.
29. Al-Hashimi HM, et al. (2000) Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J Magn Reson* 143(2):402–406.
30. 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* 40(4):489–502.
31. Chernomordik LV, Kozlov MM (2003) Protein-lipid interplay in fusion and fission of biological membranes. *Annu Rev Biochem* 72(2003):175–207.
32. Lakomek NA, et al. (2013) Internal dynamics of the homotrimeric HIV-1 viral coat protein gp41 on multiple time scales. *Angew Chem Int Ed Engl* 52(14):3911–3915.
33. 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(3):926–938.
34. Sun Z-YJ, et al. (2008) HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane. *Immunity* 28(1):52–63.
35. Reuven EM, et al. (2012) HIV-1 gp41 transmembrane domain interacts with the fusion peptide: Implication in lipid mixing and inhibition of virus-cell fusion. *Biochemistry* 51(13):2867–2878.
36. Kilby JM, Eron JJ (2003) Novel therapies based on mechanisms of HIV-1 cell entry. *N Engl J Med* 348(22):2228–2238.
37. Hildinger M, et al. (2001) Membrane-anchored peptide inhibits human immunodeficiency virus entry. *J Virol* 75(6):3038–3042.
38. 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(7):3249–3258.
39. 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(4):e60302.
40. Golding H, et al. (2002) Dissection of human immunodeficiency virus type 1 entry with neutralizing antibodies to gp41 fusion intermediates. *J Virol* 76(13):6780–6790.
41. Gustchina E, et al. (2010) Structural basis of HIV-1 neutralization by affinity matured Fabs directed against the internal trimeric coiled-coil of gp41. *PLoS Pathog* 6(11):e1001182.
42. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160(1):65–73.

Supporting Information

Roche et al. 10.1073/pnas.1401397111

SI Materials and Methods

Sample Purification. The six-helix bundle (6HB) complex formed by an ectodomain ($Core^S$) that contains large segments of the N-terminal heptad repeat (NHR) and C-terminal heptad repeat (CHR) helices connected by a six-residue linker (SGGRGG) sequence was amplified by PCR and cloned into the pET15b vector (Novagen; EMD Millipore Chemicals) between the Nde1 and BamH1 sites. Expression in *Escherichia coli* BL21(DE3) yields $Core^S$ containing the flanking residues GSSHHHHHHS SGLVPRGSHM, derived from the vector, at the N terminus of $Core^S$. Because the expressed protein is insoluble, cell lysis and initial purification of the 6H-$Core^S$ by Ni-NTA affinity chromatography were carried out in 6 M guanidine hydrochloride in 50 mM Tris·HCl, pH 8. The protein was then dialyzed against 25 mM Tris·HCl buffer, pH 7.5, containing 100 mM NaCl, 2 mM $CaCl_2$, and 20 mM imidazole, which is suitable for thrombin cleavage to remove the majority of the N-terminal nonnative residues, including the 6H-tag. The digest was reapplied on the Ni-NTA agarose column under denaturing conditions to retain the undigested 6H-$Core^S$ and the 6His-tag. The flow through containing $Core^S$ was concentrated and subjected to size-exclusion chromatography (SEC; Superdex-75; GE Healthcare) in 4 M guanidine hydrochloride, 50 mM Tris·HCl, pH 8, 5 mM EDTA, and 1 mM DTT followed by reversed-phase HPLC. The longer ectodomain construct ($Core^{IL}$) was expressed without the 6H-tag, isolated from the insoluble fraction (inclusion bodies), and subjected to SEC followed by reversed-phase HPLC as described above. Peak fractions of the protein eluting roughly in 35% (vol/vol) acetonitrile/water containing 0.05% trifluoroacetic acid were pooled and stored frozen at −70 °C. Constructs were verified both by DNA sequencing and electrospray ionization MS. Isotope-enriched proteins were grown in minimal media with the appropriate isotope source in either H_2O or D_2O.

Another plasmid construct bearing the $Core^S$ sequence flanked by the residues GSSHHHHHHSSG at the N terminus and SGLVPRGSGG residues instead of the L6 spacer was also expressed. Initial isolation was carried out as described above. After thrombin cleavage of the SGLVPRGSGG linker, the protein was denatured in 6 M guanidine hydrochloride and subjected to Ni-NTA affinity chromatography to separate the 6H-NHR (retained) from the CHR (flow through). Additional purification was carried out as described above for the $Core^S$ by SEC and HPLC. The purified 6H-NHR and CHR peptides are of the following sequences, respectively: GSSHHHHHHS SG^{546}SGIVQQQN NLLRAIEAQQ HLLQLTVWGI KQLQAR^{579}SGLV PR and GSGG^{625}HTTWME WDREINNYTS LIHSLIEESQ NQQEKNEQEL LE^{662}.

$Core^S$ and $Core^{IL}$ were folded by dialysis in 50 mM sodium formate, pH 3.0, dialyzed in 50 mM sodium acetate, pH 4, concentrated, and stored. The same dialysis scheme was also followed for the NHR and CHR peptides. The calculated molecular masses of purified $Core^S$, $Core^{IL}$, NHR, and CHR correspond to 8,284, 12,792, 4,975, and 5,791 Da, respectively.

Molecular Mass Analysis. Molecular masses were analyzed by analytical SEC with in-line multiangle light scattering (Wyatt-925-H2HC, DAWN Heleos; Wyatt Technology Inc.) with refractive index (Wyatt-215-TRXH; Wyatt Technology Inc.) and UV (Waters 2487; Waters Corporation) detectors. Volumes of injection ranged from 100 to 150 μL. Typically, 200 μg total protein were applied to a preequilibrated Superdex-75 column (1.0 × 30 cm; GE Healthcare) at a flow rate of 0.5 mL/min at room temperature and eluted in either 50 mM acetate (pH 4.0) or 20 mM sodium phosphate (pH 6.0). Guanidine hydrochloride (0.2 M) was included in the column buffer to prevent nonspecific binding of the protein to the column matrix. When studying the influence of dodecyl phosphocholine (DPC) on destabilization of $Core^S$ and $Core^{IL}$ trimers, the sample containing 10 mM DPC was layered and fractionated in the same column buffer with the inclusion of 2 mM DPC. Molecular masses were calculated using the Astra software provided with the instrument. Concentrations of proteins sampled for measurements are *ca*. 5 μM. For mass analysis, both $Core^S$ and $Core^{IL}$ contained an N-terminal sequence GSSHHHHHHS SGLVPRGS (6H), contributing to the total mass of the protein. Calculated masses of 6H-$Core^S$ and 6H-$Core^{IL}$ are 10,035 and 14,955 Da, respectively.

CD. CD spectra were recorded at 310 K in either 25 mM sodium acetate buffer at pH 4.0 or 20 mM sodium phosphate at pH 6.0 on a JASCO J-810 spectropolarimeter using a 0.1-cm path-length cell. The CD spectra of $Core^S$ and $Core^{IL}$ were recorded at a protein concentration of 15 μM, whereas 30 μM was used for the NHR and CHR peptide samples. α-Helical content was determined using the CDNN program (1).

NMR Spectroscopy. ^1H-^{15}N transverse relaxation optimized spectroscopy-heteronuclear single quantum coherence (TROSY-HSQC) spectra of $Core^S$ were recorded on a uniformly (>95%) $^2H/^{15}N/^{13}C$-enriched sample at 0.3 mM (monomer concentration) in either 50 mM sodium acetate (pH 4.0) or 20 mM sodium phosphate (pH 6.0) at 310 K in the absence and presence of 100 mM DPC using a 600 MHz Bruker Avance II spectrometer equipped with a *z* axis TCI cryogenic probe.

The backbone assignment of $Core^S$ in the absence and presence of DPC was based on 3D TROSY-HNCO and 3D TROSY-HNCACB spectra recorded at 600 MHz on the $^2H/^{15}N/^{13}C$-enriched sample at 0.3 mM in 50 mM sodium acetate (pH 4.0) at 310 K and confirmed by 3D ^{15}N-separated NOESY-HSQC spectra. Similar experiments were performed for the NHR and CHR peptides in the presence of DPC. The side chain assignments of $Core^S$ in the presence of 100 mM DPC are based on a constant time 3D H(CC)(CO)NH experiment recorded at 600 MHz on uniformly $^{15}N/^{13}C$-labeled peptide at 0.85 mM in 50 mM sodium acetate (pH 4.0) at 310 K.

The ^{15}N spin–lattice (R_1) and spin–spin ($R_{1\rho}$) and the steady state heteronuclear ^{15}N-{^1H} NOE data were collected at a ^1H frequency of 600 MHz using TROSY-based ^1H-^{15}N heteronuclear experiments (2) on uniformly $^2H/^{15}N$-labeled samples in both the absence and presence of 100 mM DPC. The R_2 rates were derived from $R_{1\rho}$ values measured with a radiofrequency spin lock of 1.3 kHz by correcting them in the standard manner (3) for ^{15}N radiofrequency offset. All relaxation experiments were performed at a protein concentration of 0.75 mM in 50 mM sodium acetate (pH 4.0) at 310 K.

The backbone residual dipolar couplings (RDCs), $^1D_{NH}$, $^1D_{NC'}$, $^2D_{HNC'}$ and $^1D_{C\alpha C'}$, were measured using uniformly $^2H/^{15}N/^{13}C$-enriched $Core^S$ in 50 mM sodium acetate (pH 4.0) at 310 K. Alignment of $Core^S$ in the absence of detergent was obtained in a 4.5% neutral stretched acrylamide gel, which was radially compressed from a 6.0-mm diameter into a 4.1-mm inner diameter NMR tube. A 5.5% acrylamide gel containing 30% of the cationic DADMAC-acrylamide copolymer, radially compressed from 6 to 4.1 mm, was used to obtain the alignment of $Core^S$ in the presence of 100 mM DPC. The $^1D_{NH}$ RDCs were derived

from the difference in $^{1}J_{NH}$ + $^{1}D_{NH}$ splitting measured at 800 MHz using an ARTSY-HSQC (amide RDCs by TROSY Spectroscopy-HSQC) experiment (4) on an isotropic sample and an aligned sample. The $^{1}D_{NC'}$ and $^{2}D_{HNC'}$ RDCs were derived from the difference in $^{1}J_{NC'}$ + $^{1}D_{NC'}$ and $^{2}J_{HNC}$ + $^{2}D_{HNC'}$ splitting, respectively, which were measured in the ^{15}N and ^{1}H dimensions of a 2D TROSY-HSQC spectrum recorded at 800 MHz in the absence of $^{13}C'$ decoupling (5). The $^{1}D_{C\alpha C'}$ RDCs were derived from the difference in $^{1}J_{C\alpha C'}$ + $^{1}D_{C\alpha C'}$ splitting measured in the $^{13}C'$ dimension of 3D TROSY-HNCO spectra recorded at 500 MHz in the absence of $^{13}C^{\alpha}$ decoupling during $^{13}C'$ evolution.

Structure Calculation. A structural ensemble of the Core^{S} monomer was obtained using the Xplor-NIH program package v. 2.34 (6) through a simulated annealing protocol, which included 10,000 steps, with the temperature linearly ramped down from 2,000 to 1 K, followed by 500 steps of Powell energy minimization. Fitted experimental restraints included 465 sequential and short-range NOE distance restraints, 114 backbone dihedral angle restraints, and 157 backbone RDCs ($^{1}D_{NH}$, $^{1}D_{C'N}$, $^{2}D_{C'HN}$, and $^{1}D_{C'Ca}$) measured in the presence of 100 mM DPC. In addition, backbone/backbone hydrogen bonding geometries were restrained through a database-derived potential of mean force (7), with force constant multipliers of 0.3 and 0.1 for the directional and linearity terms, respectively.

The RDCs were fitted using two separate alignment tensors for the NHR and CHR helices, which were defined relative to a single floating axis system, with their magnitude and rhombicity left floating during the structure calculation (8). Magnitude values converged to 14.4 Hz (NHR) and 12.3 Hz (CHR), and rhombicity values converged to 0.39 (NHR) and 0.26 (CHR); these values fall close to the values obtained when best fitting RDCs for residues Q22–A33 (NHR) and residues N49–E60 (CHR) to the corresponding helical segments of the 6HB X-ray structure (Protein Data Bank ID code 1DF4). The modest difference in alignment tensor parameters for the NHR and CHR indicates that they undergo differential motions relative to the micelle protein body, which was expected for two helices that are flexibly tethered.

1. Böhm G, Muhr R, Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng* 5(3):191–195.
2. Lakomek NA, Ying JF, Bax A (2012) Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* 53(3):209–221.
3. Peng JW, Thanabal V, Wagner G (1991) 2D Heteronuclear NMR measurements of spin-lattice relaxation—times in the rotating frame of X nuclei in heteronuclear HX spin systems. *J Magn Reson* 94(1):82–100.
4. Fitzkee NC, Bax A (2010) Facile measurement of ^{1}H-^{15}N residual dipolar couplings in larger perdeuterated proteins. *J Biomol NMR* 48(2):65–70.
5. Wang YX, et al. (1998) Simultaneous measurement of H-1-N-15, H-1-C-13′, and N-15-C-13′ dipolar couplings in a perdeuterated 30 kDa protein dissolved in a dilute liquid crystalline phase. *J Am Chem Soc* 120(29):7385–7386.
6. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160(1):65–73.
7. Grishaev A, Bax A (2004) An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J Am Chem Soc* 126(23):7281–7292.
8. Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2001) An easy way to include weak alignment constraints into NMR structure calculations. *J Biomol NMR* 21(3):275–280.

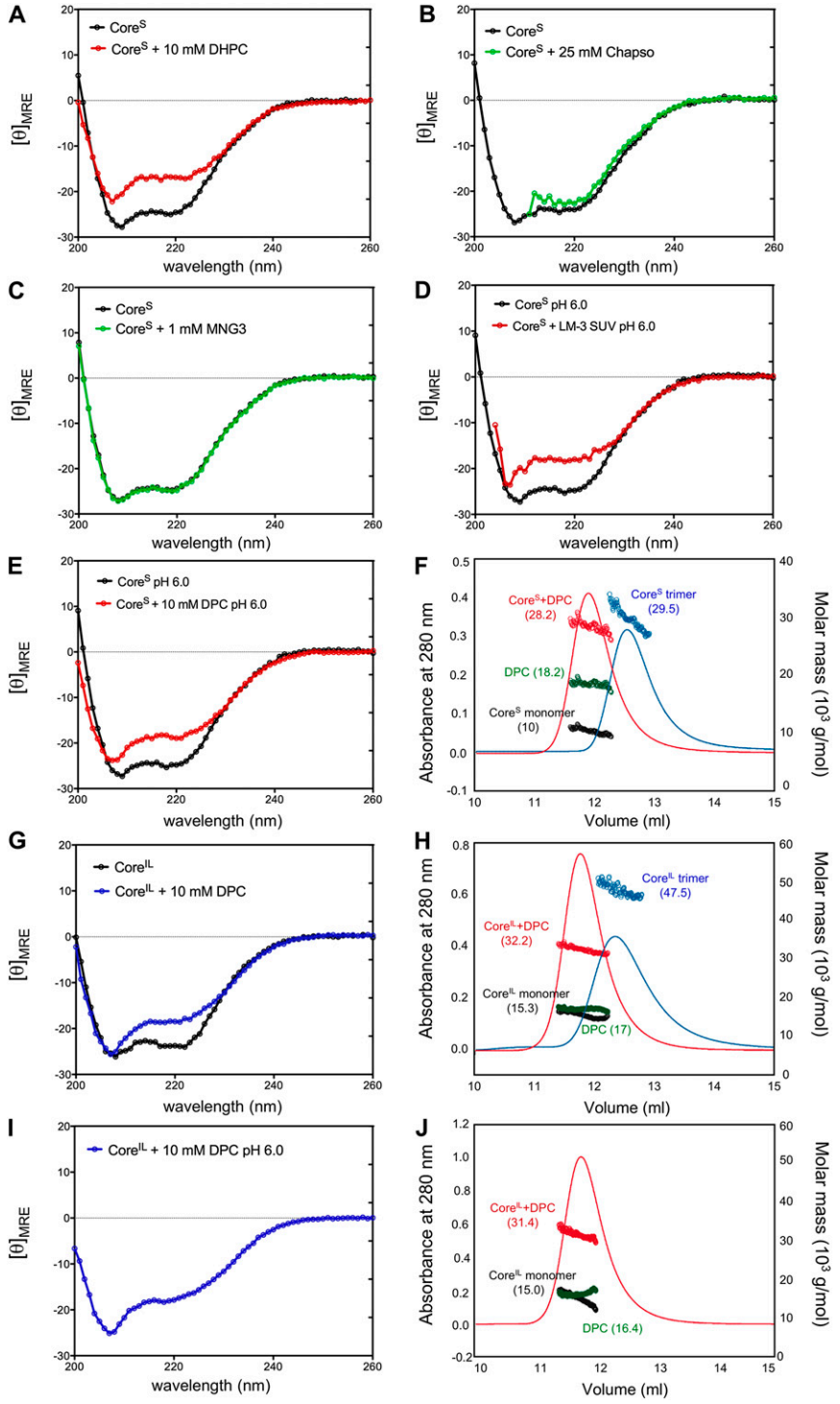

Fig. S1. CD spectra, reported here as the mean residue ellipticity (10^3 degrees centimeter^2 decimoles^{-1} residue^{-1}), of Core^S recorded at pH 4.0 (*A*–*D*) in the absence of detergent (black traces) and with (*A*) 10 mM DHPC (red trace), (*B*) 25 mM CHAPSO (green trace), and (*C*) 1 mM MNG3 (green trace). (*C* and *E*) CD spectra of Core^S recorded at pH 6.0 in the absence of detergent or vesicles (black traces), with (*D*) 5 mM LM3 vesicles (red trace) and (*E*) 10 mM DPC (red trace). (*F*) SEC with in-line multiangle light scattering (SEC-MALS) data of Core^S recorded at pH 6.0. The elution profiles monitored by the absorbance at 280 nm and masses (circles) are shown for the Core^S trimer (blue, 29.5 ± 0.3 kDa) and the complex formed by a Core^S monomer bound to a DPC micelle (red, 28.2 ± 0.2 kDa). The composition of this complex is shown in black (Core^S monomer, 10 ± 0.1 kDa) and green (DPC micelle, 18.2 ± 0.3 kDa) circles, respectively. (*G*) CD spectra of Core^{IL} recorded at pH 4.0 in the absence of detergent (black) and with 10 mM DPC (blue). (*H*) SEC-MALS data of Core^{IL} recorded at pH 4.0. The elution profiles monitored by the absorbance at 280 nm and masses (circle) are shown for the Core^{IL} trimer (blue, 47.5 ± 0.6 kDa) and the micelle-associated Core^{IL} monomer (red, 32.2 ± 0.5 kDa). The composition of this complex is shown in black (Core^{IL} monomer, 15.3 ± 0.2 kDa) and green (DPC micelle, 17 ± 0.4 kDa) circles, respectively. (*I*) CD spectra of Core^{IL} recorded at pH 6.0, in the presence of 10 mM DPC. (*J*) SEC-MALS data of Core^{IL} recorded at pH 6.0 in the presence of DPC, yielding a mass of 31.4 ± 0.2 kDa contributed by Core^{IL} (15 ± 0.1 kDa) and the DPC micelle (16.4 ± 0.3 kDa). In the absence of DPC, Core^{IL} is insoluble at pH 6.

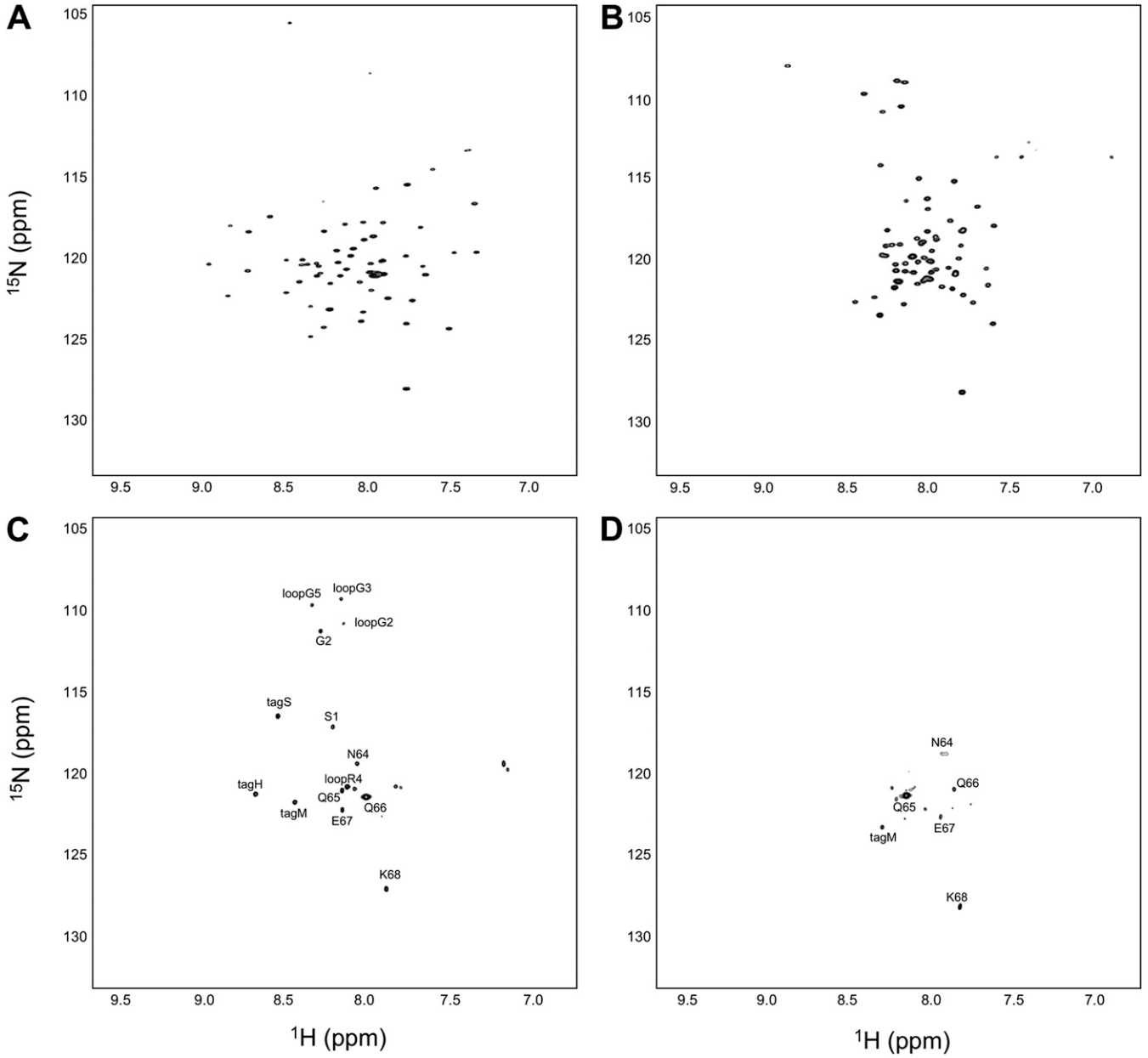

Fig. S2. ^{1}H-^{15}N TROSY-HSQC spectra of Core^{S} recorded at 310 K in 20 mM sodium phosphate (pH 6.0) in (*A*) the absence of detergent and (*B*) the presence of 100 mM DPC. Assignments of the amide resonances are listed in Table S1. ^{1}H-^{15}N TROSY-HSQC spectra of Core^{S} recorded at 310 K in the presence of 10 mM LM3 vesicles (POPC:POPE:POPS:shingomyelin:cholesterol at 10:5:2:2:10 molar ratio) at (*C*) pH 4.0 and (*D*) pH 6.0. Assignments of the amide resonances shown in *C* and *D* were derived by titrating the Core^{S}:LM3 solution with increasing concentrations of DPC. Residues labeled with tag refer to the nonnative residues at the N terminus of Core^{S} (*SI Materials and Methods*, *SI Sample Purification*), whereas loop refers to the six-residue linker connecting the NHR and CHR regions.

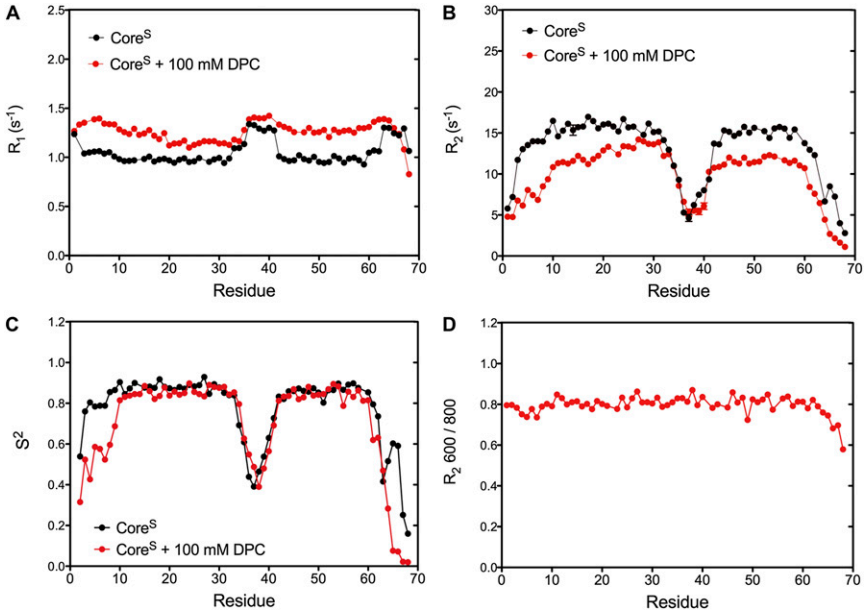

Fig. S3. Comparison of the Core^S ^{15}N relaxation properties in the absence (black) and presence (red) of 100 mM DPC detergent. (*A*) R_1. (*B*) R_2. All data were recorded at a ^{1}H frequency of 600 MHz on uniformly ^{2}H/^{15}N-enriched Core^S samples in 50 mM sodium acetate (pH 4.0) at 310 K. (*C*) Order parameters S^2 extracted from a model free analysis (1) of the experimentally determined ^{15}N relaxation parameters R_1, R_2, and heteronuclear NOE. (*D*) Ratios of the R_2 rates measured at ^{1}H frequencies of 600 and 800 MHz for Core^S in the presence of 100 mM DPC.

1. 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(17):4546–4559.

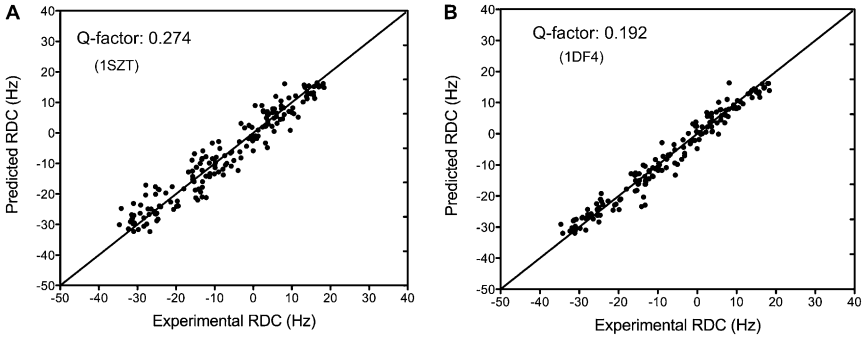

Fig. S4. Comparison of the experimental and predicted normalized RDC values extracted from a singular value decomposition fit to the 6HB crystal structures of Core^S: (*A*) 1SZT and (*B*) 1DF4. The backbone RDCs, $^1D_{NH}$, $^1D_{NC'}$, $^2D_{HNC'}$, and $^1D_{C\alpha C'}$, were measured from the difference in splitting between an isotropic sample and a sample aligned in 4.5% neutral stretched acrylamide gel radially compressed from a 6.0-mm diameter into a 4.1-mm inner diameter NMR tube. RDCs are normalized relative to ^{15}N-^{1}H values by scaling them by factors of 8.33 ($^1D_{NC'}$), 3.33 ($^2D_{HNC'}$), and 5.05 ($^1D_{C\alpha C'}$) (1). RDCs were collected in the absence of detergent using uniformly ^{2}H/^{15}N/^{13}C-enriched Core^S samples in 50 mM sodium acetate (pH 4.0) at 310 K.

1. Ottiger M, Bax A (1998) Determination of relative N-H, N-C′, Ca-C′, and Ca-Ha effective bond lengths in a protein by NMR in a dilute liquid crystalline phase. *J Am Chem Soc* 120(47): 12334–12341.

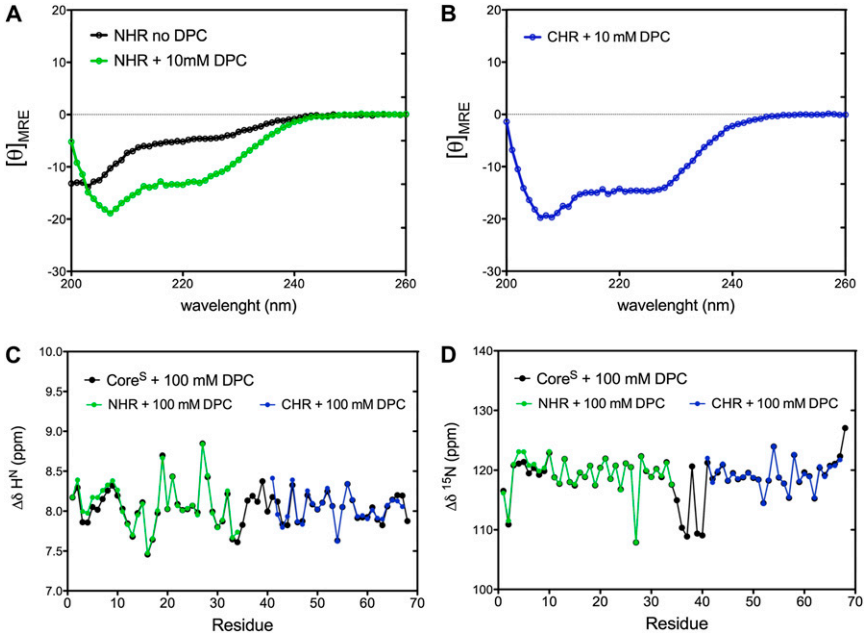

Fig. S5. CD spectra, reported as the mean residue ellipticity (10^3 degrees centimeter^2 decimoles^{-1} residue^{-1}), of the individual (*A*) NHR and (*B*) CHR peptides recorded at 30 μM peptide concentration in 50 mM sodium acetate (pH 4.0) at 310 K in (*A*) the absence of detergent (black trace) and with (*A* and *B*) 10 mM DPC (green and blue traces, respectively). In the absence of DPC, CHR is insoluble in the buffer used. (*C* and *D*) Comparison of the (*C*) $^1H^N$ chemical shifts and (*D*) ^{15}N chemical shifts measured for Core^S (black) and the individual NHR (green) and CHR (blue) peptides in the presence of 100 mM DPC. The small differences observed at the N- and C-terminal regions of the NHR and CHR regions, respectively, presumably reflect the presence of additional residues that are necessary for the purification of the NHR and CHR peptides (*SI Materials and Methods*).

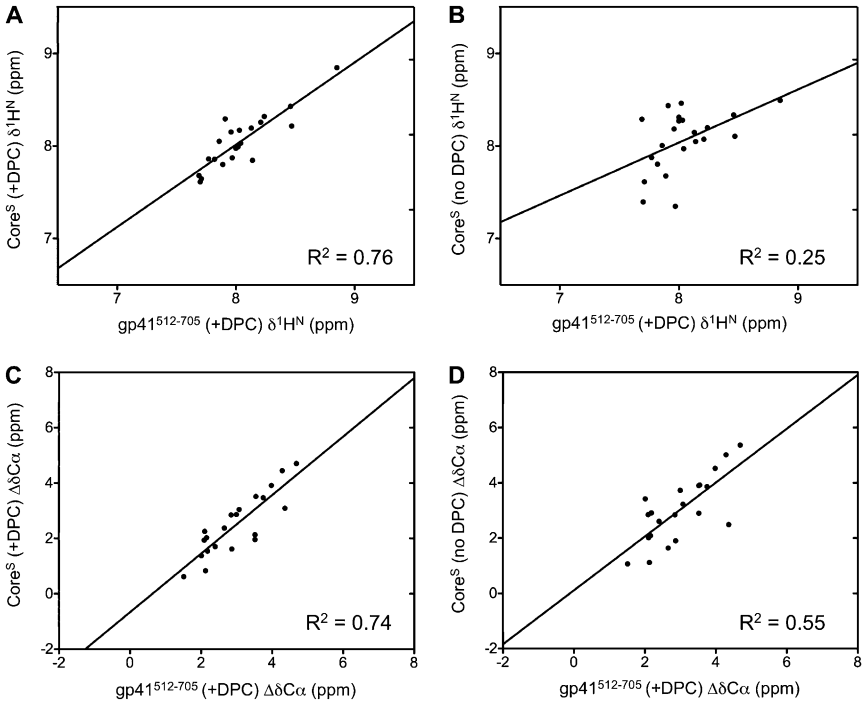

Fig. S6. Comparison of the $^1H^N$ chemical shifts reported for the gp41$^{512–705}$ in the presence of DPC (1) and measured in the present study for the Core^S (*A*) monomer and (*B*) trimer. Comparison of the secondary $^{13}C^\alpha$ chemical shifts of gp41$^{512–705}$ in the presence of DPC (1) and for the Core^S (*C*) monomer and (*D*) trimer. The comparison includes 22 ectodomain residues of gp41$^{512–705}$ for which chemical shifts were previously reported, and they are all located in the NHR helix.

1. Lakomek N-A, et al. (2013) Internal dynamics of the homotrimeric HIV-1 viral coat protein gp41 on multiple time scales. *Angew Chem Int Ed Engl* 52(14):3911–3915.

Table S1. Backbone chemical shifts of Core^S in the absence of detergent (no DPC) and the presence of 100 mM DPC (+DPC) in 50 mM sodium acetate (pH 4.0) and 25 mM sodium phosphate (pH 6.0) at 310 K

	Core^S H^N pH 4.0 (no DPC)	Core^S ^{15}N pH 4.0 (no DPC)	Core^S $^{13}C'$ pH 4.0 (no DPC)	Core^S $^{13}C^\alpha$ pH 4.0 (no DPC)	Core^S $^{13}C^\beta$ pH 4.0 (no DPC)	Core^S H^N pH 6.0 (no DPC)	Core^S ^{15}N pH 6.0 (no DPC)	Core^S H^N pH 4.0 (+DPC)	Core^S ^{15}N pH 4.0 (+DPC)	Core^S $^{13}C'$ pH 4.0 (+DPC)	Core^S $^{13}C^\alpha$ pH 4.0 (+DPC)	Core^S $^{13}C^\beta$ pH 4.0 (+DPC)	Core^S H^N pH 6.0 (+ DPC)	Core^S ^{15}N pH 6.0 (+ DPC)	Core^{IL} H^N pH 4.0 (+DPC)	Core^{IL} ^{15}N pH 4.0 (+DPC)	Core^{IL} $^{13}C'$ pH 4.0 (+DPC)	Core^{IL} $^{13}C^\beta$ pH 4.0 (+DPC)	Core^{IL} $^{13}C^\beta$ pH 4.0 (+DPC)
tags*								8.552	116.19	174.76	57.89	63.41							
tagH*								8.809	121.32	174.83	55.65	28.28							
tagM*	8.411	122.12	176.25	55.46	32.19			8.523	121.03	176.40	55.78	31.95							
S546	8.279	117.40	175.65	58.60	63.12	8.269	116.43	8.171	116.54	175.43	59.11	62.91	8.128	116.14					
G547	8.436	111.78	175.49	46.00		8.305	111.56	8.294	110.89	174.89	45.55		8.272	110.70					
I548	7.877	121.79	177.93	63.19	37.21	7.876	122.38	7.861	120.80	176.74	62.25	37.46	7.829	120.62					
V549	7.804	122.78	178.10	65.39	31.00	7.728	122.52	7.857	121.09	176.99	63.62	31.18	7.847	121.49	7.990	123.32	176.34	62.58	31.50
Q550	8.005	121.47	178.29	58.12	27.62	7.953	121.01	8.051	121.38	176.96	56.75	28.25	8.023	121.02	8.192	123.62	176.30	56.04	28.49
Q551	8.155	119.92	178.52	57.94	28.04	8.102	119.77	8.018	119.48	176.72	56.12	28.22	7.973	119.19	8.177	121.07	176.28	55.99	28.49
Q552	8.184	119.70	177.71	58.91	27.75	8.187	119.44	8.151	120.36	176.18	56.85	28.08	8.136	120.43	8.233	121.00	175.88	56.15	28.48
N553	8.073	118.94	177.46	55.70	37.88	8.087	119.34	8.256	119.23	176.30	54.80	38.20	8.254	118.90	8.285	119.91	175.72	53.99	38.36
N554	8.197	120.67	177.50	55.85	37.27	8.127	120.59	8.319	119.84	177.15	54.94	37.68	8.273	119.45	8.322	120.31	176.72	54.62	38.00
L555	8.147	120.30	178.61	57.69	40.38	8.178	120.18	8.194	122.90	178.05	57.70	40.76	8.143	122.45	8.259	123.20	178.02	57.63	40.85
L556	7.971	120.23	178.78	57.85	40.41	7.982	120.25	8.030	118.80	178.83	57.67	40.23	7.944	118.51	7.944	118.68	178.78	57.49	40.27
R557	8.049	117.81	179.94	59.02	29.58	8.027	117.71	7.845	117.72	178.84	58.62	29.05	7.789	117.99	7.785	117.81	178.67	58.42	29.07
A558	8.290	124.25	177.81	55.34	16.21	8.266	124.17	7.681	121.87	180.04	54.48	17.50	7.720	122.33	7.687	121.90	179.94	54.39	17.56
I559	8.310	120.93	179.27	64.95	37.44	8.308	121.01	7.975	118.01	177.76	64.34	37.20	7.861	117.35	7.895	117.84	177.63	64.09	37.21
E560	8.509	121.47	179.25	59.38	28.24	8.494	122.04	8.110	117.43	177.03	57.92	27.42	8.057	119.86	8.046	117.77	177.18	57.94	27.60
A561	8.080	123.75	180.20	55.23	18.02	8.039	123.79	7.457	119.58	178.57	52.65	18.08	7.640	120.27	7.501	119.85	178.57	52.79	18.11
Q562	8.972	120.35	178.16	59.76	27.97	8.964	120.29	7.642	118.84	176.16	54.63	26.38	7.779	117.92	7.636	118.61	176.26	54.94	26.90
Q563	8.425	120.17	178.64	57.12	26.03	8.401	120.33	7.975	120.73	177.48	58.29	28.12	7.977	120.50	8.003	120.79	177.46	58.32	28.00
H564	7.436	118.55	177.64	58.75	26.33	7.474	119.58	8.697	117.47	176.33	56.96	27.42			8.587	117.36	176.26	57.07	27.52
L565	7.607	120.49	178.51	58.01	41.46	7.663	120.41	8.026	120.41	179.68	57.74	40.83	7.945	120.32	7.907	120.30	179.43	57.51	40.90
L566	8.811	122.05	180.31	57.86	40.82	8.848	122.23	8.433	121.93	178.54	58.13	40.13	8.324	122.02	8.344	121.68	178.38	57.88	40.27
Q567	8.283	120.30	179.52	59.30	27.06	8.296	120.41	8.087	118.54	179.43	58.90	27.47	8.038	118.73	8.037	118.40	179.01	58.54	27.52
L568	7.954	121.57	182.13	58.25	40.73	7.977	121.89	8.013	120.75	178.42	57.63	41.29	8.004	120.86	7.920	120.61	178.34	57.37	41.39
T569	8.871	117.85	176.63	66.88	68.09	8.834	117.92	8.024	116.80	176.18			7.997	116.63	8.016	116.13	176.07	66.83	67.82
V570	8.322	124.67	177.71	67.71	30.90	8.346	124.74	8.068	121.13	177.60	66.82	30.83	8.060	121.20	8.061	120.99	177.48	66.48	30.84
W571	8.271	120.10	178.84	62.09	28.40	8.311	120.23	7.983	120.52	178.70	61.43	27.74	7.975	119.84	7.896	120.54	178.47	61.01	27.91
G572	8.493	105.38	174.87	47.00		8.473	105.45	8.846	107.88	174.99	47.24		8.850	107.90	8.727	107.65	175.04	46.98	
I573	8.337	122.69	177.74	65.43	37.42	8.347	122.88	8.427	122.33	178.03	64.86	36.51	8.440	122.31	8.429	122.06	177.84	64.49	36.40
K574	8.462	119.80	179.73	59.68	30.96	8.493	120.03	7.993	119.85	180.05	59.27	30.68	7.985	119.81	7.960	120.34	179.99	59.24	30.74
Q575	7.676	118.05	178.90	57.22	27.23	7.678	118.02	7.800	118.87	178.75	57.16	26.86	7.796	118.87	7.799	118.62	178.62	57.16	26.94
L576	7.350	119.41	179.66	56.48	41.33	7.338	119.55	7.873	120.22	179.30	57.21	40.66	7.870	120.23	7.951	120.47	178.37	57.10	40.92
Q577	8.105	120.89	177.87	57.62	26.78	8.165	121.00	8.216	118.83	177.91	58.22	28.02	8.217	118.84	8.277	118.92	177.44	59.50	27.83
A578	7.613	121.09	178.88	53.74	17.31	7.647	120.93	7.647	121.34	179.05	53.46	17.72	7.629	121.28	7.696	119.64	180.45	54.44	17.42
R579	7.396	116.64	176.90	56.52	29.43	7.349	116.56	7.612	117.58	176.80	56.23	29.65	7.594	117.66	7.639	117.86	177.94	57.06	29.21
loopS1†	7.617	114.37	174.83	58.16	63.42	7.604	114.45	7.827	114.96	175.29	58.93	63.37	7.836	114.94					
loopG2†	7.988	110.34	174.50	45.00				8.132	110.34	174.84	45.34		8.159	110.38					
loopG3†	8.033	108.56	174.48	44.96		7.985	108.55	8.190	108.88	174.40	44.97		8.185	108.81					
loopR4†	8.373	119.96	177.49	56.41	29.49			8.118	120.63	176.90	56.05	29.75	8.083	120.51					
loopG5†	8.604	110.60	175.37	46.13				8.374	109.37	174.69	45.12		8.386	109.60					

Table S1. Cont.

	$Core^S$ H^N pH 4.0 (no DPC)	$Core^S$ ^{15}N pH 4.0 (no DPC)	$Core^S$ $^{13}C'$ pH 4.0 (no DPC)	$Core^S$ $^{13}C^\alpha$ pH 4.0 (no DPC)	$Core^S$ $^{13}C^\beta$ pH 4.0 (no DPC)	$Core^S$ H^N pH 6.0 (no DPC)	$Core^S$ ^{15}N pH 6.0 (no DPC)	$Core^S$ H^N pH 4.0 (+DPC)	$Core^S$ ^{15}N pH 4.0 (+DPC)	$Core^S$ $^{13}C'$ pH 4.0 (+DPC)	$Core^S$ $^{13}C^\alpha$ pH 4.0 (+DPC)	$Core^S$ $^{13}C^\beta$ pH 4.0 (+DPC)	$Core^S$ H^N pH 6.0 (+ DPC)	$Core^S$ ^{15}N pH 6.0 (+ DPC)	$Core^{IL}$ H^N pH 4.0 (+DPC)	$Core^{IL}$ ^{15}N pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C'$ pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C^\beta$ pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C^\beta$ pH 4.0 (+DPC)
loopG6[†]	8.398	109.66	175.54	45.39				7.995	109.06	174.68	45.09		8.137	108.91					
I580															7.973	120.22	177.47	64.24	36.67
L581															8.038	120.55	179.46	57.61	40.20
A582															7.532	121.48	180.53	54.60	17.55
V583															7.879	118.86	177.51	65.73	30.65
E584															8.323	119.46	177.95	59.61	27.28
R585															7.917	118.48	178.02	58.83	29.02
Y586															7.738	120.07	177.51	60.57	37.94
L587															8.087	117.75	178.47	56.72	40.96
K588															7.802	118.66	178.12	58.05	31.50
D589															8.041	119.73	176.94	55.08	39.38
Q590															7.840	118.39	176.12	55.67	27.80
Q591															7.996	118.14	176.77	56.56	27.24
L592															7.992	120.40	177.30	55.90	41.46
L593															7.889	118.37	178.30	55.90	41.02
G594															8.029	107.96	175.28	45.68	
I595															7.659	119.85	176.47	62.51	37.38
W596															7.915	121.55	177.30	57.91	28.93
G597															8.126	108.18	174.65	45.55	
A598															7.912	123.65	178.45	52.81	18.35
S599															8.048	114.37	175.46	59.67	62.93
G600															8.142	110.25	174.56	45.80	
K601															7.768	120.33	177.18	56.80	31.87
L602															7.910	121.41	177.54	56.03	41.39
I603															7.742	118.89	174.75	61.36	37.60
A604																			
T605																			
T606															7.712	115.37	173.90	61.94	69.13
A607															7.998	125.77	177.54	51.94	18.72
V608																			
P609																			
W610															7.590	119.71	175.56	57.30	28.99
N611															7.621	118.94	174.77	52.67	38.18
A612															7.849	124.15	177.44	52.50	18.34
S613															7.979	114.03	174.61	58.44	63.11
W614															7.898	122.55	175.86	56.79	28.87
S615															7.914	115.37	174.55	58.02	63.53
N616															8.293	121.01	175.93	54.30	38.18
K617															8.097	121.43	177.62	57.30	31.75
S618															7.971	115.99	175.94	59.86	62.80
L619															8.144	122.38	178.30	57.14	40.60
E620															8.050	117.88	177.96	58.69	27.76
Q621															7.718	117.39	178.17	57.80	27.91

Table S1. Cont.

	$Core^S$ H^N pH 4.0 (no DPC)	$Core^S$ ^{15}N pH 4.0 (no DPC)	$Core^S$ $^{13}C'$ pH 4.0 (no DPC)	$Core^S$ $^{13}C^\alpha$ pH 4.0 (no DPC)	$Core^S$ $^{13}C^\beta$ pH 4.0 (no DPC)	$Core^S$ H^N pH 6.0 (no DPC)	$Core^S$ ^{15}N pH 6.0 (no DPC)	$Core^S$ H^N pH 4.0 (+DPC)	$Core^S$ ^{15}N pH 4.0 (+DPC)	$Core^S$ $^{13}C'$ pH 4.0 (+DPC)	$Core^S$ $^{13}C^\alpha$ pH 4.0 (+DPC)	$Core^S$ $^{13}C^\beta$ pH 4.0 (+DPC)	$Core^S$ H^N pH 6.0 (+ DPC)	$Core^S$ ^{15}N pH 6.0 (+ DPC)	$Core^{IL}$ H^N pH 4.0 (+DPC)	$Core^{IL}$ ^{15}N pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C'$ pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C^\beta$ pH 4.0 (+DPC)	$Core^{IL}$ $^{13}C^\beta$ pH 4.0 (+DPC)
I622															7.650	118.69	172.24	63.64	37.40
W623															7.948	117.78	176.09	59.11	
N624																			
H625																		56.10	
T626															7.658	114.97	175.52	62.91	69.17
T627																			
W628	7.963	120.97	176.76	57.61	28.69	7.938	120.98	8.176	121.26	177.30	58.58	28.69	8.188	121.04	8.251	122.07	176.01	59.13	28.54
M629	8.047	121.18	178.99	57.60	31.48	8.047	121.38	8.120	118.57	177.93	57.46	30.92	8.063	118.43	8.021	119.20	178.20	57.03	27.78
E630	8.536	121.08	177.92	58.69	27.30	8.728	120.69	7.834	119.60	177.98	57.52	27.56	7.808	119.66	7.707	119.49	178.24	57.79	27.83
W631	7.803	123.39	177.21	62.03	28.63	7.764	123.94	7.824	120.89	177.48	58.96	28.42	7.827	120.53	7.871	120.85	177.51	59.36	28.60
D632	8.730	117.94	178.74	57.15	40.32	8.722	118.29	8.325	118.21	177.07	55.95	38.26	8.194	120.02	8.391	118.27	177.22	56.10	38.49
R633	7.774	120.00	179.06	59.20	29.31	7.767	119.78	7.860	119.52	178.41	58.44	29.19	8.019	119.60	7.831	119.21	178.50	58.54	29.19
E634	8.102	120.67	179.04	59.20	28.53	7.967	120.99	7.874	118.47	178.17	57.81	27.72	8.086	119.51	7.815	118.41	178.29	57.88	27.70
I635	8.423	120.53	180.16	63.11	34.88	8.366	120.30	8.199	118.80	178.06	64.03	36.99	8.253	119.48	8.250	118.66	178.22	64.04	36.91
N636	8.178	121.54	178.74	56.10	37.25	8.226	121.47	8.085	119.58	177.11	55.70	38.08	8.134	119.96	8.043	119.54	177.16	55.84	38.15
N637	8.300	120.87	178.44	56.00	37.30	8.286	120.84	8.020	118.69	177.12	55.06	37.68	8.029	118.67	7.995	118.71	177.15	55.15	37.74
Y638	8.268	118.01	178.03	63.04	37.37	8.264	118.26	8.106	118.43	177.48	60.43	37.71	8.166	118.80	8.077	118.25	177.49	60.48	37.75
T639	8.012	114.52	176.59			8.593	117.36	8.247	114.47	176.97	66.22	67.58	8.286	113.97	8.226	114.41	176.95	66.21	67.62
S640	7.949	118.31	176.96	61.44	62.03	7.906	117.73	8.064	118.26	176.99	61.30	62.02	7.999	118.00	8.031	118.18	176.96	61.32	62.04
L641	7.574	124.27	179.13	57.91	40.98	7.505	124.25	7.631	123.95	179.06	57.69	41.00	7.599	123.63	7.600	123.87	179.04	57.70	41.04
I642	8.420	119.70	177.40	65.77	37.47	8.395	120.01	8.051	118.77	177.58	64.52	36.46	7.950	118.33	8.006	118.63	177.57	64.46	36.47
H643	8.123	117.72	177.13	59.22	27.34	8.137	117.83	8.339	117.77	176.80	59.09	27.60	8.243	117.93	8.298	117.70	176.76	59.10	27.62
S644	7.979	115.62	176.25	61.56	62.22	7.949	115.61	8.137	115.35	175.99	61.58	62.28	8.051	114.77	8.087	115.22	175.90	61.53	62.35
L645	8.070	123.34	180.13	57.44	41.68	8.026	123.23	7.914	122.54	178.94	57.29	41.19	7.781	121.88	7.861	122.31	178.80	57.24	41.27
I646	8.419	120.54	177.97	64.88	36.83	8.414	121.37	7.921	118.01	178.12	63.83	37.15	7.694	116.50	7.845	117.62	177.93	63.63	37.21
E647	7.900	120.43	178.81	58.53	27.86	7.938	120.82	7.925	119.64	177.95	57.56	27.71	7.911	121.37	7.884	119.79	177.81	57.50	27.92
E648	8.040	118.74	178.29	57.79	27.90	8.021	118.78	8.048	118.99	177.45	57.10	27.62	8.190	120.39	8.016	119.05	177.33	57.02	27.77
S649	7.765	115.64	175.41	60.54	63.28	7.759	115.39	7.893	115.26	174.94	59.62	63.14	8.000	116.00	7.874	115.28	174.87	59.49	63.17
Q650	7.984	120.79	175.66	56.67	28.24	7.901	120.89	7.824	120.40	176.07	55.80	28.70	7.983	120.93	7.819	120.52	176.00	55.80	28.70
N651	7.935	120.08	176.77	53.62	38.22	7.908	120.07	8.059	119.18	175.30	53.28	38.34	8.096	119.55	8.037	119.20	175.25	53.27	38.37
Q652	7.997	118.51	175.97	55.79	29.22	7.964	118.57	8.144	120.74	175.89	55.70	28.53	8.173	121.05	8.102	120.72	175.84	55.71	28.56
Q653	7.980	120.91	175.92	55.75	28.32	7.988	120.78	8.199	121.07	175.84	55.61	28.67	8.200	121.43	8.152	121.05	175.80	55.64	28.73
E654	8.199	122.58	175.23	55.91	28.82	8.229	123.06	8.194	122.32	175.20	55.66	28.41	8.289	123.11	8.155	122.34	175.15	55.75	28.57
K655	7.793	127.34		57.08	32.76	7.762	127.95	7.875	127.05		56.76	32.56	7.788	127.81	7.795	127.14		56.87	32.66

The backbone assignment of $Core^{IL}$ (C607A/C612A) in the presence of 100 mM DPC in 50 mM sodium acetate (pH 4.0) at 318 K is also reported. Chemical shifts are in parts per million relative to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). Residue numbering follows the numbering of full-length gp41.

*The three residues named tagS, tagH, and tagM correspond to an uncleaved fragment of the original His-tag, which is present only in the $Core^S$ construct (*SI Materials and Methods*).

†The six residues replacing the immunodominant loop (I69–T116) in the $Core^S$ construct are numbered loopS1–loopG6.

Table S2. Backbone dynamics and partitioning of the Core^S monomer at the water–micelle interface

	R_1 (s^{-1}) 600 MHz	R_2 (s^{-1}) 600 MHz	NOE 600 MHz	R_1 (s^{-1}) 800 MHz	R_2 (s^{-1}) 800 MHz	S^2	τ_e (ps)	Omniscan (2 mM)	16-DSA (2 mM)
S546	1.27 (±0.03)	4.80 (±0.04)	0.03 (±0.01)	1.11 (±0.04)	6.03 (±0.01)			0.43	0.62
G547	1.33 (±0.04)	4.75 (±0.27)	0.17 (±0.02)	1.17 (±0.07)	5.96 (±0.01)	0.32 (±0.01]	272 (±9)	0.38	0.75
I548	1.35 (±0.02)	6.76 (±0.01)	0.27 (±0.01)	1.14 (±0.03)	8.63 (±0.01)	0.52 (±0.02)	695 (±2)	0.64	0.67
V549		6.16 (±0.07)	0.27 (±0.01)	1.01 (±0.09)	8.19 (±0.05)	0.43 (±0.02)	986 (±10)	0.58	0.69
Q550	1.39 (±0.05)	8.05 (±0.02)	0.29 (±0.01)	1.17 (±0.07)	10.90 (±0.03)	0.59 (±0.03)	697 (±7)	0.60	0.64
Q551	1.40 (±0.06)	7.42 (±0.02)	0.33 (±0.01)	1.15 (±0.10)	9.54 (±0.03)	0.58 (±0.04)	724 (±8)	0.56	0.75
Q552	1.34 (±0.06)	6.85 (±0.03)	0.41 (±0.01)	1.10 (±0.10)	9.30 (±0.03)	0.52 (±0.03)	707 (±7)	0.51	0.56
N553	1.34 (±0.06)	8.48 (±0.03)	0.46 (±0.01)	1.09 (±0.08)	10.75 (±0.02)	0.60 (±0.05)	657 (±10)	0.50	0.70
N554	1.33 (±0.07)	9.36 (±0.04)	0.56 (±0.02)	1.05 (±0.09)	11.67 (±0.04)	0.69 (±0.04)	631 (±14)	0.52	0.76
L555	1.29 (±0.06)	10.84 (±0.05)	0.64 (±0.01)	1.01 (±0.08)	13.70 (±0.05)	0.82 (±0.03)	76 (±4)	0.65	0.50
L556	1.26 (±0.02)	11.33 (±0.01)	0.67 (±0.02)	0.95 (±0.02)	13.36 (±0.02)	0.83 (±0.03)	55 (±1)		
R557	1.24 (±0.03)	11.46 (±0.01)	0.67 (±0.01)	0.95 (±0.04)	13.79 (±0.03)	0.84 (±0.05)	49 (±2)	0.67	0.82
A558	1.29 (±0.05)	11.27 (±0.04)	0.68 (±0.01)	1.00 (±0.06)	14.08 (±0.04)	0.85 (±0.04)	64 (±3)	0.72	0.75
I559	1.23 (±0.05)	11.59 (±0.03)	0.68 (±0.01)	0.92 (±0.06)	14.30 (±0.05)	0.85 (±0.04)	46 (±3)	0.73	0.66
E560	1.24 (±0.06)	12.21 (±0.03)	0.70 (±0.01)	0.95 (±0.06)	14.97 (±0.05)	0.89 (±0.03)	36 (±5)	0.74	0.58
A561	1.28 (±0.08)	11.73 (±0.08)	0.66 (±0.02)	0.98 (±0.11)	14.85 (±0.07)	0.86 (±0.03)	67 (±6)	0.70	0.50
Q562	1.22 (±0.07)	11.21 (±0.05)	0.64 (±0.01)	0.92 (±0.08)	13.97 (±0.06)	0.82 (±0.08)	52 (±4)	0.68	0.49
Q563	1.19 (±0.09)	11.81 (±0.10)	0.70 (±0.01)	0.86 (±0.07)	15.20 (±0.07)	0.84 (±0.04)	30 (±5)	0.46	0.52
H564	1.25 (±0.09)	12.22 (±0.07)	0.68 (±0.02)	0.93 (±0.08)	14.96 (±0.08)	0.88 (±0.03)	42 (±8)	0.32	0.39
L565	1.12 (±0.05)	12.88 (±0.07)	0.70 (±0.02)	0.83 (±0.05)	16.06 (±0.07)	0.84 (±0.04)	0 (±4)	0.62	0.44
L566	1.14 (± 0.04)	13.33 (±0.06)	0.72 (±0.01)	0.84 (±0.04)	16.82 (±0.06)	0.86 (±0.05)	0 (±3)	0.68	0.37
Q567	1.14 (±0.06)		0.72 (±0.01)	0.78 (±0.06)	16.62 (±0.09)	0.84 (±0.05)	12 (±4)	0.67	0.67
L568	1.17 (±0.07)	12.40 (±0.08)	0.72 (±0.02)	0.85 (±0.06)	15.94 (±0.07)	0.85 (±0.04)	13 (±5)	0.66	0.54
T569	1.10 (±0.05)	13.39 (±0.04)	0.74 (±0.01)	0.78 (±0.04)	16.06 (±0.06)	0.90 (±0.04)	0 (±13)	0.75	0.29
V570	1.13 (±0.05)	13.35 (±0.06)	0.74 (±0.01)	0.83 (±0.04)	16.99 (±0.05)	0.86 (±0.03)	0 (±4)	0.74	0.43
W571	1.14 (±0.04)	13.15 (±0.07)	0.71 (±0.01)	0.85 (±0.05)	15.85 (±0.05)	0.85 (±0.03)	8 (±3)	0.70	0.71
G572	1.16 (± 0.10)	14.21 (±0.29)	0.75 (±0.02)	0.85 (±0.10)	16.47 (±0.09)	0.83 (±0.08)	12 (±6)	0.77	0.44
I573	1.16 (±0.06)	13.99 (±0.08)	0.75 (±0.01)	0.87 (±0.06)	17.27 (±0.06)	0.89 (±0.04)	0 (±6)	0.76	0.43
K574	1.17 (±0.05)	13.67 (±0.06)	0.76 (±0.01)	0.84 (±0.04)	16.87 (±0.06)	0.88 (±0.03)	0 (±4)	0.76	0.77
Q575	1.14 (±0.07)	13.62 (±0.08)	0.73 (±0.01)	0.84 (±0.06)	16.94 (±0.07)	0.88 (±0.04)	0 (±6)	0.76	0.81
L576	1.14 (±0.08)	13.86 (±0.11)	0.74 (±0.02)	0.83 (±0.08)	16.64 (±0.09)	0.88 (±0.05)	0 (±7)	0.77	0.66
Q577	1.13 (±0.10)	12.20 (±0.08)	0.69 (±0.01)	0.87 (±0.07)	15.47 (±0.07)	0.84 (±0.05)	25 (±4)	0.76	0.64
A578	1.18 (±0.07)	12.42 (±0.08)	0.67 (±0.02)	0.90 (±0.08)	15.59 (±0.07)	0.85 (±0.04)	28 (±5)	0.70	0.78
R579	1.16 (±0.05)	10.95 (±0.02)	0.58 (±0.01)	0.90 (±0.05)	13.51 (±0.07)	0.80 (±0.01)	52 (±2)	0.67	0.71
loopS1	1.28 (±0.04)	8.56 (±0.06)	0.51 (±0.01)	1.01 (±0.05)	10.32 (±0.04)	0.63 (±0.01)	96 (±2)	0.57	0.65
loopG2	1.39 (±0.05)	6.61 (±0.31)	0.45 (±0.01)	1.14 (±0.07)	7.96 (±0.03)	0.55 (±0.02)	385 (±23)	0.45	0.52
loopG3	1.41 (±0.03)	5.29 (±0.42)	0.36 (±0.01)	1.22 (±0.06)	6.51 (±0.02)	0.49 (±0.05)	469 (±14)	0.32	0.66
loopR4	1.40 (±0.03)	5.52 (±0.10)	0.37 (±0.01)	1.19 (±0.05)	6.34 (±0.01)	0.39 (±0.02)	817 (±3)	0.37	0.67
loopG5	1.40 (±0.03)	5.43 (±0.38)	0.35 (±0.01)	1.21 (±0.04)	6.83 (±0.01)	0.48 (±0.03)	433 (±11)	0.31	0.64
loopG6	1.42 (±0.06)	6.07 (±0.40)	0.46 (±0.02)	1.24 (±0.13)	7.26 (±0.03)	0.57 (±0.01)	578 (±26)	0.36	0.49
W628		10.27 (±0.12)	0.53 (±0.02)	1.10 (±0.21)	11.51 (±0.06)	0.69 (±0.13)	1,290 (±99)		
M629	1.33 (±0.07)	10.76 (±0.04)	0.64 (±0.02)	1.01 (±0.07)	13.75 (±0.05)	0.81 (±0.03)	77 (±4)	0.57	0.55
E630	1.31 (±0.04)	10.88 (±0.04)	0.64 (±0.01)	0.98 (±0.04)	13.59 (±0.03)	0.83 (±0.02)	81 (±3)	0.62	0.65
W631	1.29 (±0.09)	11.14 (±0.07)	0.65 (±0.01)			0.84 (±0.04)	75 (±6)	0.69	0.59
D632	1.26 (±0.09)	12.00 (±0.08)	0.67 (±0.01)	0.98 (±0.11)	15.30 (±0.08)	0.87 (±0.05)	53 (±8)	0.70	0.48
R633		11.48 (±0.09)	0.64 (±0.01)	0.97 (±0.07)	13.36 (±0.07)	0.82 (±0.06)	49 (±3)	0.63	0.77
E634	1.25 (±0.06)	11.27 (±0.04)	0.64 (±0.01)	0.95 (±0.06)	13.93 (±0.05)	0.83 (±0.02)	63 (±3)	0.68	0.70
I635	1.30 (±0.11)	11.99 (±0.07)	0.70 (±0.02)	0.96 (±0.08)	14.39 (±0.06)	0.88 (±0.05)	59 (±7)	0.73	0.65
N636	1.25 (±0.06)	11.27 (±0.06)	0.69 (±0.01)	0.93 (±0.05)	15.57 (±0.06)	0.84 (±0.03)	50 (±4)	0.65	0.60
N637	1.26 (±0.09)	11.45 (±0.07)	0.67 (±0.01)	1.00 (±0.09)	13.88 (±0.06)	0.84 (±0.04)	55 (±6)		
Y638	1.28 (±0.18)	11.52 (±0.06)	0.67 (±0.02)	0.91 (±0.08)	14.21 (±0.07)	0.85 (±0.04)	52 (±5)	0.75	0.50
T639	1.21 (±0.05)	12.11 (±0.04)	0.70 (±0.01)	0.90 (±0.06)	14.70 (±0.06)	0.87 (±0.03)	21 (±4)	0.70	0.41
S640	1.28 (±0.08)	12.29 (±0.06)	0.72 (±0.01)	0.96 (±0.07)	14.50 (±0.06)	0.90 (±0.03)	46 (±7)	0.64	0.62
L641	1.26 (±0.06)	12.13 (±0.04)	0.71 (±0.01)	0.94 (±0.04)	15.67 (±0.06)	0.88 (±0.02)	36 (±6)	0.65	0.65
I642	1.27 (±0.08)		0.67 (±0.01)	0.98 (±0.05)	13.05 (±0.05)	0.79 (±0.04)	37 (±1)	0.73	0.69
H643	1.28 (±0.04)	11.67 (±0.02)	0.70 (±0.01)	0.98 (±0.05)	14.08 (±0.04)	0.86 (±0.01)	52 (±3)	0.72	0.66
S644	1.25 (±0.06)	11.35 (±0.02)	0.69 (±0.02)	0.94 (±0.07)	13.53 (±0.05)	0.83 (±0.01)	49 (±4)	0.67	0.74
L645	1.30 (±0.07)	11.62 (±0.06)	0.68 (±0.01)	0.98 (±0.07)	14.68 (±0.06)	0.86 (±0.04)	66 (6)	0.70	0.70
I646	1.30 (±0.04)	11.04 (±0.01)	0.68 (±0.01)	0.98 (±0.04)	13.57 (±0.04)	0.81 (±0.02)	66 (±2)	0.76	0.71
E647	1.31 (±0.05)	10.71 (±0.04)	0.65 (±0.01)	1.00 (±0.05)	13.18 (±0.04)	0.82 (±0.04)	77 (±3)	0.69	0.73
E648	1.36 (±0.05)	8.42 (±0.04)	0.38 (±0.02)	1.09 (±0.05)	10.78 (±0.06)	0.62 (±0.05)	701 (±8)	0.64	0.77
S649	1.39 (±0.05)	7.59 (±0.06)	0.45 (±0.0.02)	1.12 (±0.07)	9.23 (±0.02)	0.63 (±0.06)	694 (±10)	0.65	0.68

Table S2. Cont.

	R_1 (s^{-1}) 600 MHz	R_2 (s^{-1}) 600 MHz	NOE 600 MHz	R_1 (s^{-1}) 800 MHz	R_2 (s^{-1}) 800 MHz	S^2	τ_e (ps)	Omniscan (2 mM)	16-DSA (2 mM)
Q650	1.39 (±0.05)	6.45 (±0.02)	0.33 (±0.01)	1.17 (±0.07)	8.16 (±0.02)	0.47 (±0.02)	749 (±6)	0.61	0.73
N651	1.38 (±0.05)	4.43 (±0.01)	0.32 (±0.01)	1.16 (±0.06)	5.83 (±0.01)	0.28 (±0.05)	873 (±4)	0.56	0.74
Q652	1.30 (±0.03)	2.69 (±0.01)	−0.14 (±0.01)	1.20 (±0.04)	3.61 (±0.01)	0.08 (±0.03)	687 (±2)	0.46	0.73
Q653	1.24 (±0.03)	2.14 (±0.02)	−0.37 (±0.01)	1.16 (±0.04)	3.14 (±0.02)	0.07 (±0.04)	623 (±1)	0.48	0.75
E654	1.08 (±0.02)	1.63 (±0.02)	−0.71 (±0.01)	1.06 (±0.03)	2.33 (±0.01)	0.02 (±0.04)	482 (±1)	0.48	0.75
K655	0.83 (±0.02)	1.11 (±0.08)	−1.24 (±0.02)	0.83 (±0.02)	1.92 (±0.01)	0.02 (±0.04)	112 (±1)	0.52	0.78

The ^{15}N relaxation rates R_1 and R_2, measured at 600 and 800 MHz, are reported as well as the heteronuclear NOE (measured at 600 MHz), the order parameter S^2, and the internal motion correlation times (τ_e) derived from a Modelfree analysis (1) of the ^{15}N spin relaxation data recorded for the Core^S monomer. The attenuation ratios induced by the paramagnetic agents gadodiamide (Omniscan) and 16-doxyl-stearic acid (16-DSA) on the amide cross-peak intensities of Core^S are also reported. All data were recorded in 50 mM sodium acetate (pH 4.0) at 310 K with 100 mM DPC.

1. 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(17):4546–4559.

Table S3. Structural statistics for the 20 lowest energy models of the Core^S monomer

rmsd from experimental restraints (no. of restraints)	Value
NOE (465)	0.076 (0.01)
Dihedral (114)	5.7° (0.1)
RDC tensor parameters (no. of RDCs)	
Da(NH)/R, NHR	14.4 Hz/0.39
Da(NH)/R, CHR	12.3 Hz/0.26
Fitted $^1D^{NH}$ Q factor (22) NHR	9.6%
Fitted $^1D^{C\alpha H\alpha}$ Q factor (22) NHR	8.7%
Fitted $^1D^{C'N}$ Q factor (20) NHR	10.3%
Fitted $^2D^{C'HN}$ Q factor (20) NHR	9.9%
Fitted $^1D^{NH}$ Q factor (21) CHR	9.1%
Fitted $^1D^{C\alpha H\alpha}$ Q factor (20) CHR	8.3%
Fitted $^1D^{C'N}$ Q factor (17) CHR	11.4%
Fitted $^2D^{C'HN}$ Q factor (17) CHR	10.8%
Average database H bond directional energy	−3.9 kT
Average database H bond linearity energy	0.49 kT
Ramachandran statistics	
ϕ/ψ in most favored regions	98.5%
ϕ/ψ in additionally allowed regions	1.5%
Atomic rmsd	
Backbone heavy atoms (residues 4–64)	0.251 Å