Published in final edited form as:
Biochim Biophys Acta. 2009 September ; 1788(9): 1804–1812. doi:10.1016/j.bbamem.2009.06.011.

Molecular Dynamics Studies of the Transmembrane Domain of Gp41 from HIV-1

Jong Hwa Kim[a,*], **Taryn L. Hartley**[a], **A. Rachael Curran**[b], and **Donald M. Engelman**[b]

[a] Department of Chemistry, Alabama A&M University, Normal, AL 35762

[b] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT06520

Abstract

Helix-helix interactions in the putative three-helix bundle formation of the gp41 transmembrane (TM) domain may contribute to the process of virus-cell membrane fusion in HIV-1 infection. In this study, molecular dynamics is used to analyze and compare the conformations of monomeric and trimeric forms of the TM domain in various solvent systems over the course of 4 to 23-ns simulations. The trimeric bundles of the TM domain were stable as helices and remained associated in a hydrated POPE lipid bilayer for the duration of the 23-ns simulation. Several stable inter-chain hydrogen bonds, mostly among the three deprotonated arginine residues located at the center of each of the three TM domains, formed in a right-handed bundle embedded in the lipid bilayer. No such bonds were observed when the bundle was left-handed or when the central arginine residue in each of the three TM helices was replaced with isoleucine (R_I mutant), suggesting that the central arginine residues may play an essential role in maintaining the integrity of the three-helix bundle. These observations suggest that formation of the three-helix bundle of the TM domain may play a role in the trimerization of gp41, thought to occur during the virus-cell membrane fusion process.

Keywords

HIV-1; gp41; transmembrane (TM); molecular dynamics

1. Introduction

Extensive studies of the human immunodeficiency virus type 1 (HIV-1) over the past two decades have given us a great deal of information regarding the function of each of its viral components and the mechanism of cell infection in causing disease [1]. HIV-1 consists of one RNA and 15 proteins. Among these proteins, two envelope glycoproteins, gp120 and gp41, produced from proteolytic cleavage of a precursor peptide, gp160 [2], are responsible for initiating cell infection. It is widely accepted that binding of the surface envelope protein (gp120) of the mature virion to receptors of the cell is followed by subsequent virus-cell membrane fusion by the transmembrane envelope protein (gp41), leading to entry of the core into the cell [3–5].

*Corresponding author. Department of Chemistry, Alabama A&M University, Normal, AL 35762, USA. Tel: 1-256-372-4935; fax: 1-256-372-8288. E-mail address: jong.kim@aamu.edu.

The three-dimensional structure of the entire gp41 molecule (345 amino acids) is unknown, but the envelope glycoprotein is generally believed to be a trimeric complex, with its six individual subunits (three gp120 and three gp41 subunits) held together by noncovalent interactions [6–8]. Two crystal structures of a small core fragment of the gp41 protein have been reported [9,10]. The core region is a six-helix bundle consisting of the interior parallel coiled-coil trimer of N36 and the exterior antiparallel C34, which is similar to the low-pH induced conformation of influenza hemagglutinin; thus, the core region is probably not the resting structure of gp41 but rather a structure formed during the fusion reaction. Proposed models for cell fusion by gp41 have so far been based on the structure of the core region [3, 4], and most current studies of gp41 in the cell fusion process are focused on it. In confirmation of the proposed models, small peptides similar in structure to C34 or N36 of the core region have been found to inhibit the fusion process. A new generation of anti-HIV drugs based on this information is in clinical trials by pharmaceutical companies [11,12].

Two significant regions of gp41 are not included in the core region: an N-terminal hydrophobic glycine-rich “fusion” peptide (FP), which is believed to initiate fusion, and the transmembrane (TM) domain, which has an important role in anchoring the envelope proteins in the viral membrane and in fusion [13]. Recently, the structure of the gp41 fusion peptide has been studied by using molecular dynamics simulations [14] and by spectroscopic and electron microscopic methods [15–17]. Since TM domains are essential components of structural integrity for many membrane proteins [18], it is interesting to consider whether the TM domain of gp41 may be involved in oligomerization of the protein in the native environment. Mutation studies suggest that the gp41 TM domain is important for the membrane fusion process [19], and the gp41 TM region is thought to play an important role in anchoring the envelope proteins in the viral membrane during HIV-cell fusion [13]. Our TOXCAT (a measure of transmembrane helix association in a biological membrane [20]) study using the 22 amino acid-TM domain (^{685}IFIMIVGGLVGLRIVFAVLSIV^{706}) of gp41 (and its mutated fragments) from HIV-1 ARV2/SF2 isolate suggests that the TM domain associates. (See Results and discussion.)

Despite its possible role in the structure and function of gp41, the difficulty of studying it has resulted in little attention being paid to the TM domain of gp41, and most research is focused on the ectodomain (the core region). Insights on the structure of the TM domain and its self-association would be useful in understanding how the TM domain contributes to the overall oligomeric conformation of gp41 and to the virus-cell membrane fusion process.

It has been difficult to obtain atomic-level structures of single TM integral membrane proteins because these proteins do not crystallize well. However, computational methods have proven to be very useful for investigation of the TM domains of membrane proteins, and have been successful in suggesting structures that later proved correct (*e.g.*, glycophorin [21,22]). In this paper, we report a molecular dynamics (MD) study of monomers and possible three-helix bundles of the gp41 TM domain in a hydrated lipid bilayer and other solvent systems. Stability in different environments and the structures of the bundles are studied in order to better understand gp41 TM domain oligomerization.

2. Materials and methods

2.1. TOXCAT

TOXCAT constructs consist of an N-terminal DNA-binding domain of ToxR, a transmembrane domain, and the periplasmic maltose-binding protein. ToxR′ (gp41-TM)MBP chimerae contained the complete TM domain, with the amino sequence of ^{685}IFIMIVGGLVGLRIVFAVLSIV^{706}. Individual constructs were cloned by mutating the wild-type gp41 construct using the QuikChange site-directed mutagenesis kit (Stratagene)

protocol using oligonucleotides ~20 base pairs long. Constructs were transformed into *E. coli* NT326 (MalE-) cells. Wholecell lysates were used to estimate expression levels of the constructs. Samples were run on SDS-PAGE, and then Western analysis was carried out using antibodies against MBP (NEB). Blots were developed using goat anti-rabbit alkaline phosphatase secondary antibody (Pierce). The presence of MBP in the periplasm was confirmed by growth on minimal maltose media.

CAT Assays—Cell-free extracts were made by pelleting 200 μl of cells at an A_{600} of 0.6, resuspending in 500 μl of 0.1M Tris, pH 8.0, then lysing with 20 μl of 100 mM EDTA, 100 mM dithiothreitol, and 50 mM Tris, pH 8.0, and one drop of toluene, at 30 °C for 30 min. The cell-free extract was then diluted 1:80 before being used in the CAT assays. Briefly, samples were incubated at 37 °C with tritiated chloramphenicol and n-butyryl coenzyme A. After 90 min, the reaction was halted by partitioning the [^{3}H]chloramphenicol-butyryl CoA complex into xylene. The organic phase was washed and quantified using the radiolabel. All measurements were performed four times. Errors shown are standard deviations on four measurements.

2.2. Determination of the initial structures

The 20-residue TM domain (^{1}FIMIVGGLVGLRIVFAVLSI^{20}) of gp41 from the HIV-1 ARV2/SF2 isolate was used in the study. Starting structures for trimeric helix bundles of gp41 TM domains *in vacuo* were determined at the Center for Structural Biology Computation Lab at Yale using a global searching program, CHI (CNS Searching of Helix Interactions [21]), which utilizes the molecular dynamics program CNS (Crystallography & NMR system [23]) with the OPLS force field (optimized potentials for liquid simulations, Jorgensen and Tirado-Rives [24]). This method has proven to be successful in predicting the structures of transmembrane proteins [25,26]. Three parallel α-helices were positioned with the distance between centers of neighboring helices set at 0.4 Å and an initial crossing angle at 25 °. A set of structures were obtained by rotating the helices from 0 ° to 360 ° for each starting structure (both left-handed and right-handed) using a 45 ° step size. For each structure, four trials were run using simulated annealing of all atomic coordinates, with rotation and crossing angles free to vary (36×4×2=288 structures for symmetric searches, where the helices were rotated cooperatively to preserve a threefold axis; 8×8×8×4×2=4096 structures for full searches without the threefold constraint on the starting structures). Root mean square deviations (RMSDs) of atomic coordinates between structures were then determined to select clusters of similar structures, thought to represent basins of convergence. An average structure for each cluster and the spread of structures in each cluster were found, and the interaction energy between residues of the averaged structure for each cluster was calculated. Structures with high interaction energy and large contact area were chosen as starting structures for the molecular dynamics simulations.

2.3. Molecular dynamics

Molecular dynamics simulations of gp41 TM single helix strands and trimeric helix bundles were performed using the MPI version of GROMACS 3.2.1 [27] running on the Cray XD1 at the Alabama Supercomputer Center in Huntsville, AL (typically running in parallel on 4 Opteron processors). For a single-helix strand or a helix bundle embedded in a hydrated lipid bilayer, the helix (bundle) was inserted in a hydrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylethanolamine (POPE) lipid bilayer (using the topology and structure for the lipid bilayer from Tieleman and Berendsen [28]). The single helix (or the helix bundle) was placed on the lipid bilayer by generating a box of the solvent, namely the hydrated POPE, around the helix (or the helix bundle). The helix (or the helix bundle) was inserted in the center of the box in a perpendicular orientation to the lipid bilayer plane. Solvent (lipid and water) molecules were removed from the box when the distance between any atom of the helix (or the helix

bundle) and any atom of the solvent molecule was less than the sum of the van der Waals radii of both atoms. After energy minimization, the system was equilibrated for an extended period of time by using position-restrained molecular dynamics (with the positions of the helix bundle restrained but the positions of others free to move) until the density (or volume), pressure, and temperature became stabilized (typically for about 3 ns). Na^+ and Cl^- ions were included to maintain the neutral charge of the system and a physiological salt concentration of 0.154 M. The sodium and chloride ions replaced water molecules at the position of the first atoms with the most favorable electrostatic potential. In order to simulate the physiological salt concentration of 0.154 M, about one NaCl pair for every 360 water molecules or one ion per 180 water molecules was added. Since between 4000 and 6000 water molecules were typically used in each system, between 11 and 17 NaCl pairs (11~17 Na^+ and 11~17 Cl^-) were added. (Salt ions were added to water only, not to the lipid bilayer.) The density of the system was around 1.02 g/ml while the pressure and the temperature were maintained at 1.0 bar and 300 K, respectively. Full molecular dynamics was carried out using a periodic boundary condition and the isothermal-isobaric (NPT) ensemble with a time step of 1 fs (300K, 1 bar, Nosé-Hoover temperature coupling [29,30] with a time constant of 0.5 ps for each of the five groups, helix, lipid, water, Na^+, and Cl^-, and Parinello-Rahman anisotropic pressure coupling [31,32] with a coupling constant of 6.0 ps for each of the 6 components). Long-range interactions were treated with the particle-mesh Ewald method (PME [33]; interpolation order of 4 with a grid spacing of 0.12 nm) and the SPC model was used for water molecules [34].

3. Results and discussion

3.1. TOXCAT study

The level of association in the TOXCAT assay (as measured by expression of CAT) of the gp41 contruct is lower than that of glycophorin A (GpA), however, it is higher than the non-associating G83I control as seen in Fig. 1. Upon mutation of two glycine residues in the TM domain to isoleucine (G691 and G695, or residues numbers 6 and 10 in our model for the MD study), the level association decreases slightly. This is likely attributable to the role of the two glycine residues in the inter-chain interaction in association of the gp41 TM domains. Upon replacement of the arginine-697 residue (the arginine-12 residue in our model) with isoleucine, there is a similar decrease in association, implying the role of the arginine residue in oligomerization of the gp41 TM domains. Although the TOXCAT study is inconclusive as to the position of the interface, it is certain that the constructs do oligomerize in TOXCAT.

3.2. Stability of the monomeric helix of the TM domain in lipid and in water

The MD simulations of the single chain 20-residue TM domain (^{1}FIMIVGGLVGLRIVFAVLSI^{20}) were carried out in aqueous solution and in the hydrated lipid bilayer in order to examine helix stability in different environments. First, RMSDs of the C_α atoms from the initial structure over time were compared. As can be seen in Fig. 2a, the C_α RMSDs rise almost continuously to reach around 0.6 nm at the end of the 4-ns simulation for the structure in aqueous solution, implying that the structure drifts away continuously from the starting helical structure. However, in a 4 ns simulation, the C_α RMSDs fluctuate no more than 0.1 nm in the lipid bilayer, suggesting that the structure stays very close to the original helical structure in this environment.

In order to examine the fluctuations in structure, the residue-by-residue root mean square fluctuation (RMSF) of the C_α atoms from their average positions for the duration of the simulation was examined. Fig. 2b shows that the C_α RMSF values are significantly higher in water than in the lipid bilayer: among the non-terminal residues, Leu-8 and Val-9 have particularly large RMSFs in water. As can be seen in the figure, the C_α RMSF vs. the residue number in the lipid bilayer shows a relatively flat line, with even the C_α RMSF values of the

terminal residues falling below 0.1 nm. The C_α RMSF comparison suggests that the TM helix is unstable in water but very stable in the lipid environment. Other helical parameters, such as the helix radius, rise per residue, and twist per residue seen over the first 4-ns of the MD simulation (Fig. 2c,d,e) also show that the conformation of the TM domain undergoes a significant change in water, especially during the first 2 ns. However, these parameters show marked stability in the lipid bilayer (throughout the total 12 ns of simulation, although only the first 4 ns are shown in the figure). Loss of α-helicity in water is evident in the first 6–7 residues of the amino terminus (and, to a much lesser degree, in the carboxyl terminus). This can be seen from the sudden decrease in the twist angle of the helix at around 2 ns of the simulation period, as shown in Fig. 2e. Simulations using a 28-residue sequence containing the TM domain (YIKIFIMIVGGLVGLRIVFAVLSIVNRV) produce similar results. (C_α RMSD versus time for the 28-residue monomeric helix of gp41 TM domain in lipid is shown in Fig 2a.)

3.3. Conformational stability of the three-helix bundle

Conformational stabilities of helix bundles consisting of three 20-residue TM domains (^{1}FIMIVGGLVGLRIVFAVLSI^{20}) were also examined in various environments: in a hydrated lipid bilayer, in DMSO, in decane, and in water, using 5–23 ns MD simulations. The C_α RMSD vs. time and C_α RMSF analyses (Fig. 3a,b) show that by far the smallest deviations in C_α positions and RMS fluctuations are observed in the lipid bilayer, suggesting stability of the C_α positions in the helix bundle. The RMSD values in the lipid bilayer, less than 0.2 nm at the end of the 5-ns simulation, are as small as those observed in other peptide simulations in lipid bilayer environments. These include MD studies of the OmpF porin trimer [28], where the x-ray structure was used as the starting structure, and MD studies of the Influenza A M2 Channel [35], where the starting structure was obtained by MD/simulated annealing *in vacuo*. The RMSF is significantly larger in decane than in the other solvent systems, implying a very unstable conformation of the helix bundle in decane. The fact that the helix bundle shows higher stability in the lipid bilayer than in decane or DMSO suggests that the interactions between some residues in the TM domain and the POPE lipid molecules stabilize the bundle. Based on helical parameters computed in the course of the simulation period (*e.g.*, the helix radius, rise per residue, and twist per residue), each of the three helix-chains in the bundle appears to retain an α-helical conformation in the lipid bilayer during the entire simulation period, which is also supported by Ramachandran plot analysis (not shown).

In order to examine which handedness of the three-helix bundle is favored, MD simulations of right- and left-handed bundles (with each chain being right-handed α-helical) were carried out in the hydrated lipid bilayer. Fig. 3c and 3d show the C_α RMSD and the C_α RMSF plots for the right-handed and the left-handed three-helix bundles. Our data suggest that the right-handed bundle has smaller RMS fluctuations than the left-handed helix bundles (0.188±0.028 versus 0.213±0.022 nm). Unsurprisingly, the terminal residues of each helix exhibit larger RMS fluctuations for both the right- and the left-handed helix bundles. These larger fluctuations are probably due to the larger number of possible interactions of the terminal residues with water molecules and lipid head groups [35], and the lack of constraint from the parts of the structure that are not represented in the model. Analyses of the helical parameters (data not shown) suggest that the TM domains in left-handed helix bundle also remain α-helical in the course of the simulation periods. Thus, the data suggest that both the right- and left-handed helix bundles appear to be stable in the lipid bilayer for the simulation time frame used in the current study (18 to 23 ns), with the left-handed bundle somewhat less stable.

The secondary structures of the right- and left-handed three-helix bundles of the TM domains were compared using the φ and Ψ angles for each of the 18 residues in each helical chain (a total of 54 residues in the helix bundle). There appear to be no significant differences in these

angles between the right- and left-handed bundles for the duration of the simulation time frames used in the current study.

There is, however, a significant difference in terms of inter-chain hydrogen bonding between the right- and left-handed bundles. No stable inter-chain hydrogen bonds are observed in the left-handed helix bundle: only a few transient hydrogen bonds appear in the course of the 18-ns simulation. In the right-handed bundle, on the other hand, at least several stable inter-chain hydrogen bonds are present. Although the left-handed helix bundle demonstrated reasonable helical stability during the simulation time frame of 18 ns, the three helical chains might not be able to maintain a stable helix bundle in this fashion over an extended time frame in the absence of stable hydrogen bonds among the three chains. (In our simulations *in vacuo*, the right-handed helix bundle maintained its stability for the entire duration of the 50-ns simulation, but the left-handed helix bundle did not.) Further, conserved amino acids are involved in the hydrogen bonds. Therefore, we conjecture that the three gp41-molecules associate through formation of the right-handed three-helix bundle of the TM domains during cell-virus membrane fusion.

3.4. Inter-chain hydrogen bonds in the right-handed three-helix bundle

Based on the definition of Ravishanker et al. [36], the conformation of the right-handed bundle in the lipid bilayer allows several stable inter-chain hydrogen bonds. Forming within 1 or 1.5 ns of the simulation, these hydrogen bonds remain stable throughout the entire simulation period of 23 ns, as shown in Fig. 4. Most of these inter-chain hydrogen bonds form among the *deprotonated* arginine-12 (denoted Argn-12) residues located near the middle of each of the three helices (Table 1). While the Argn-12 residues for each of the three helices in the left-handed bundle are on the opposite side of the interface, these arginine residues are at the interface in the right-handed helix bundles. Therefore, these arginine residues may be essential in formation of inter-chain hydrogen bonds in the three-helix bundle. These inter-chain hydrogen bonds could play a major role in maintaining the integrity of the helix bundle over a longer time scale.

3.5. Bundle with protonated central arginine residues

Arginine residues of a protein are typically protonated in aqueous solution at physiological *pH* (the pK_R of free arginine is 12.5). Since it is of interest to see whether the central arginine residues in the three TM domains can be protonated in the helix bundle, the helical stability of the bundle with *protonated* arginine-12 (Arg-12) residues was examined with a 10-ns simulation for each of the right- and the left-handed bundles. Our MD studies show that the C_α RMSD values (Fig. 5a) of the bundle with the Arg-12 residues are larger than those for the bundle with the Argn-12 residues, and the RMSD values of the left-handed bundles are larger than those of the right-handed bundles, whether the arginine-12 residues are protonated or deprotonated. (The averaged RMSD values for the right- and the left-handed bundles during the initial 5 ns are 0.250±0.432 and 0.318±0.044 nm for the bundles with *protonated* arginine-12, and 0.188±0.028 and 0.213±0.022 nm for bundles with *deprotonated* arginine-12, respectively.) The C_α RMSF (Fig. 5b) also increases, especially in the terminal residues, upon protonation of the arginine-12 residues. Although no significant differences can be observed for helical properties in terms of helix radius, rise per residue, and twist per residue, there are larger fluctuations in these helical parameters in the bundle with the Arg-12 residues, suggesting a lower helical stability of the bundle upon protonation of these arginine residues. The Ramachandran plot shows that some residues have larger Ψ angles (non-α-helical) in comparison to the bundle with the Argn-12 residues. As the simulation progresses, there appears to be a population increase of the residues exhibiting the larger Ψ angles, which are similar to those found in β-strands. This aspect of conformation, i.e., increase of Ψ angles in many residues, resembles the bundle conformation with Argn-12 residues observed in non-

lipid environments, such as water, decane, and DMSO. These observations suggest that the helix bundle with the *deprotonated* arginine-12 residues is more stable and therefore is more likely to be a bundle in the three-helix formation of gp41.

3.6. Bundle with central arginine residues replaced with isoleucine residues (R_I mutant)

In order to determine the role of the Argn-12 residues in bundle formation, a new helical bundle was created in which each of the three arginine residues was replaced by isoleucine residues (denoted the R_I mutant) using the procedure described in Materials and methods. The helix bundle of the R_I mutants in the lipid bilayer (both right- and left-handed) were simulated for 15 ns. The helix bundle of the R_I mutants in water (both right- and left-handed) were simulated for 5 ns. The C_α RMSDs (Fig. 5c) are larger for the left-handed bundle than for the right-handed bundle (0.200±0.026 versus 0.125±0.015nm during the initial 5-ns), and the C_α RMSF plots (Fig. 5d) for both the right- and left-handed bundles of the R_I mutant are similar to those of the wild type bundles, with larger fluctuations on the helical termini. Although the Ramachandran plot of the right-handed bundle of R_I is similar to that of the wild type counterpart, some residues have larger Ψ values. In addition, residues in the left-handed bundle of the R_I mutant show a wider range of φ angles. Based on the C_α RMSD, C_α RMSF, and Ramachandran plots, the right-handed bundle of R_I appears to be rather more stable and to maintain more α-helical characteristics than the left-handed counterpart, but appears less stable than the wild type.

In water, the conformations of the helix bundles (especially the left-handed) of the R_I mutant are unstable, as expected. While the right-handed bundle shows a C_α RMSD of ~0.4 nm and a C_α RMSF of between 0.15 and 0.4 nm (except for the Phe-1 residue of Chain 3, which fluctuates up to 0.8 nm), the left-handed bundle is extremely unstable in water, with C_α RMSD reaching 2 nm and the C_α RMSF ranging between 1.1 and 1.8 nm. While the Ramachandran plots of the R_I bundles are similar to the wild type bundles in water, some residues exhibit larger Ψ angles with β-strand characters.

No stable inter-chain hydrogen bonds were observed, however, in the three-helix bundle (either right- or left-handed) of the R_I mutant in the lipid bilayer during the course of 15 ns simulations. As discussed previously, at least several stable inter-chain hydrogen bonds were observed in the right-handed helix bundle of the wild type TM domains. The fact that no stable inter-chain hydrogen bonds could be formed when the three Argn-12 residues in the bundle were replaced with the isoleucine residues emphasizes that these arginine residues may play an essential role in formation of the inter-chain hydrogen bonds in the bundle, thereby maintaining the integrity of the three-helix bundle of the gp41 TM domains.

3.7. Lengths and tilts of the helices and bundles

The average values of the lengths and tilts of the helices and the bundles during the entire simulation periods of the three-helix bundles for the wild type (both with deprotonated and protonated arginine residues) and the R_I mutant were calculated and are shown in Table 2. The tilt and the length of each of these helices and bundles fluctuate around the average values listed in the table during the entire period of the simulations; no differences in these values have been observed during the initial few ns of the simulations. The right-handed helix bundle of the wild type TM domains is slightly tilted with respect to the lipid bilayer normal, with a tilt angle of 9.9 ± 3.4 °. There are no significant differences in these tilt angles among the various helix bundles, varying between 3 and 10 °. The extent of the tilt angles of the individual helical chains in a bundle can give information regarding the degree of interactions among the chains. The tilt angles of the individual helical chains for the right-handed bundle of the wild type TM domains with respect to the lipid bilayer normal are somewhat smaller than those in the left-handed bundle of the wild type TM domains or those in the helix bundles (both right-

and left-handed) where the arginine residues in all the three chains are replaced with isoleucine residues. The tilt angles of the individual helical strands seem to become considerably larger for the helix bundles when the Argn-12 residues are protonated (for both right- and left-handed helix bundles): larger crossing angles among the helical strands in these bundles result in considerably shorter bundle lengths (around 2.55 nm compared to 2.7–2.8 nm for other helix bundles), as can be seen in Table 2. It appears that the inter-chain interactions in the bundle with protonated arginine residues gradually diminish as the simulation progresses, leading to an increase in the crossing angles and decrease in bundle lengths. Fig. 6 shows that three deprotonated arginine residues from each of the three helical chains are interacting with one another at the interfaces of the bundle.

3.8. Stability comparison of various helix bundles

In order to compare the stability of the various helix bundles discussed above, the potential energy for each of the following six types of three-helix bundles was calculated: the right- and left-handed wild-type helix bundles with deprotonated arginine (Argn-12) residues, the right- and left-handed wild-type helix bundles with protonated arginine (Arg-12) residues, and the right- and left-handed helix bundles in which arginine residues were replaced by isoleucine residues (R_I). The structures of these helix bundles were obtained from at least 10-ns simulations in the hydrated lipid bilayer. The potential energy was calculated for each helix bundle alone, by removing water, lipids, and ions from the system. Our data show that the right-handed bundle is always more stable than its left-handed counterpart for each type. In addition, the R_I mutant helix bundles are significantly less stable than the wild-type helix bundles (by about 360 kcal/mol), verifying the importance of the central arginine residue in the TM domain for stability of the helix bundle formation.

The data also show that the helix bundles with Argn-12 are considerably more stable than the helix bundles with Arg-12 (by at least 230 kcal/mol), suggesting that the central arginine residues are deprotonated in the helix bundle. The gp41 TM domain consists of all the amino acids of non-polar side chains except for the arginine residue, which is located in the middle of the TM domain. Considering the non-polar environment of the lipid bilayer, the arginine residue should favor the deprotonated (neutral) side chain. The significance of the deprotonated arginine residue, however, appears to be its essential role in formation of the three-helix bundle. Only the right-handed helix bundle appear to allow the inter-chain hydrogen bonds among these three Argn-12 residues, which stabilize the helix bundle, as verified by potential energy comparison. As a matter of fact, the right-handed wild-type helix bundle with Argn-12 is the only helix bundle that stays intact over a prolonged simulation: while all other helix bundles undergo significant deformation in various ways, the right-handed wild-type helix bundle with Argn-12 maintains its stability for the entire duration of the 50-ns simulation via its inter-chain hydrogen bonds.

4. Conclusions

Based on molecular dynamics simulations, the helical conformation of the gp41 TM domain appears to be stable in a hydrated lipid bilayer. In water, however, the simulated helix forms a kink around the Gly-7 residue, which leads to unwinding of the helix at the first several residues of the amino terminus.

The three-helix bundle of the HIV-1 gp41 TM domains shows much greater stability in the lipid bilayer than in other environments. Several stable inter-chain hydrogen bonds were observed in the right-handed helix bundle during the course of simulations, but no such bonds were observed in the left-handed helix bundle. Most of the stable inter-chain hydrogen bonds observed were formed among the three conserved deprotonated arginine residues, each located in the middle of the three TM domains. Therefore, the helix bundle is believed to be right-

handed during three-helix formation of HIV-1 gp41. Upon protonation of the central arginine residues, the helix bundle became unstable. When these arginine residues were replaced by three isoleucine residues, no such inter-chain hydrogen bonding was observed, suggesting that these arginine residues may play important roles in maintaining the stability of the helix bundle. TOXCAT data clearly show that the constructs do oligomerize, suggesting the importance of oligomerization of the TM domain. In light of the fact that stable inter-chain hydrogen bonds form within ~1 ns and remain stable throughout the entire period of the 23-ns simulations, it is very likely that formation of the right-handed three-helix bundle of the TM domain occurs (and perhaps is essential) in the formation of gp41 trimers, thought to occur during the virus-cell membrane fusion process.

Acknowledgments

We thank the National Institutes of Health for supporting this work by a grant to J.H.K. (GM071350). We also thank the Alabama Supercomputer Center for providing a grant of high performance computing resources and technical support.

References

1. Frankel AD, Young JAT. HIV-1: fifteen proteins and an RNA. Annu Rev Biochem 1998;67:1–25. [PubMed: 9759480]
2. Freed EO, Martin MA. The role of the human immunodeficiency virus type 1 envelope glycoproteins in virus infection. J Biol Chem 1995;270:23883–23886. [PubMed: 7592573]
3. Chan DC, Kim PS. HIV entry and its inhibition. Cell 1998;93:681–684. [PubMed: 9630213]
4. Doms RW, Moore JP. HIV-1 membrane fusion: targets of opportunity. J Cell Biol 2000;151:F9–F13. [PubMed: 11038194]
5. Lorizate M, Gomara MJ, de la Torre BG, Andreu D, Nieva JL. Membrane-transferring sequences of the HIV-1 gp41 ectodomain assemble into an immunogenic complex. J Mol Biol 2006;360:45–55. [PubMed: 16813835]
6. Wyatt R, Sodroski J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science 1998;280:1884–1888. [PubMed: 9632381]
7. Eckert DM, Kim PS. Mechanisms of viral membrane fusion and its inhibition. Annu Rev Biochem 2001;70:777–810. [PubMed: 11395423]
8. Poignard P, Saphire EO, Parren PW, Burton DR. Gp120: biologic aspects of structural features. Annu Rev Immunol 2001;19:253–274. [PubMed: 11244037]
9. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. Cell 1997;89:263–273. [PubMed: 9108481]
10. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. Atomic structure of the ectodomain from HIV-1 gp41. Nature 1997;387:426–430. [PubMed: 9163431]
11. Moore JP, Stevenson M. New targets for inhibitors of HIV-1 replication. Nat Rev Mol Cell Biol 2000;1:40–49. [PubMed: 11413488]
12. Cohen J. Confronting the limits of success. Science 2002;296:2320–2324. [PubMed: 12089422]
13. Hernandez LD, Hoffman LR, Wolfsberg TG, White JM. Virus-cell and cell-cell fusion. Annu Rev Cell Dev Biol 1996;12:627–661. [PubMed: 8970739]
14. Wong TC. Membrane structure of the human immunodeficiency virus gp41 fusion peptide by molecular dynamics simulation II. The glycine mutants. Biochim Biophys Acta 2003;1609:45–54. [PubMed: 12507757]
15. Martin I, Defrise-Quertain F, Decroly E, Vandenbranden M, Brasseur R, Ruysschaert JM. Orientation and structure of the NH_2-terminal HIV-1 gp41 peptide in fused and aggregated liposomes. Biochim Biophys Acta 1993;1145:124–33. [PubMed: 8422404]
16. Martin I, Schaal H, Scheid A, Ruysschaert JM. Lipid membrane fusion induced by the human immunodeficiency virus type 1 gp41 N-terminal extremity is determined by its orientation in the lipid bilayer. J Virol 1996;70:298–304. [PubMed: 8523539]

17. Gordon LM, Nisthal A, Lee AB, Eskandari S, Ruchala P, Jung CL, Waring AJ, Mobley PW. Structural and functional properties of peptides based on the N-terminus of HIV-1 gp41 and the C-terminus of the amyloid-beta protein. Biochim Biophys Acta 2008;1778:2127–2137. [PubMed: 18515070]
18. Popot JL, Engelman DM. Helical membrane protein folding, stability, and evolution. Annu Rev Biochem 2000;69:881–922. [PubMed: 10966478]
19. Helseth E, Olshevsky U, Gabuzda D, Ardman B, Haseltine W, Sodroski J. Changes in the transmembrane region of the Human Immunodeficiency virus type 1 gp41 envelope glycoprotein affect membrane fusion. J Virol 1990;64:6314–6318. [PubMed: 2243396]
20. Russ WP, Engelman DM. TOXCAT: a measure of transmembrane helix association in a biological membrane. Proc Natl Acad Sci USA 1999;96:863–868. [PubMed: 9927659]
21. Adams PD, Engelman DM, Brunger AT. An improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. Proteins 1996;26:257–261. [PubMed: 8953647]
22. MacKenzie KR, Prestegard JH, Engelman DM. A transmembrane helix dimer: structure and implications. Science 1997;276:131–133. [PubMed: 9082985]
23. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determinator. Acta Crystallogr D Biol Crystallogr 1998;54:905–921. [PubMed: 9757107]
24. Jorgensen W, Tirado-Rives J. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 1988;110:1657–1666.
25. Fleming KG, Engelman DM. Computation and mutagenesis suggest a right-handed structure for the synaptobrevin transmembrane dimmer. Proteins 2001;45:313–317. [PubMed: 11746678]
26. Leeds JA, Boyd D, Huber DR, Sonoda GK, Luu HT, Engelman DM, Beckwith JJ. Genetic selection for and molecular dynamic modeling of a protein transmembrane domain multimerization motif from a random Escherichia coli genomic library. J Mol Biol 2001;313:181–195. [PubMed: 11601855]
27. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Mod 2001;7:306–317.
28. Tieleman DP, Berendsen HJC. A molecular dynamics study of the pores formed by E. coli OmpF porin in a fully hydrated palmitoyloleoylphosphatidylethanolamine bilayer. Biophys J 1998;74:2786–2801. [PubMed: 9635733]
29. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 1984;52:255–268.
30. Hoover WG. Canonical dynamics: equilibrium phase-space distributions. Phys Rev 1985;A31:1695–1697.
31. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new Molecular dynamics method. J Appl Phys 1981;52:7182–7190.
32. Nosé S, Klein ML. Constant pressure molecular dynamics for molecular systems. Mol Phys 1983;50:1055–1076.
33. Darden T, York D, Pederson L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys 1993;98:10089–10092.
34. Berendsen, HJC.; Postma, JPM.; van Gunsteren, WF.; Hermans, J. Interaction models for water in relation to protein hydration. In: Pullman, B., editor. Intermolecular Forces, Reidel, Dordrecht. 1981. p. 331-342.
35. Forrest LR, Kukol A, Arkin IT, Tieleman DP, Sansom MSP. Exploring models of the influenza A M2 channel: MD simulations in phospholipid bilayer. Biophys J 2000;78:55–69. [PubMed: 10620273]
36. Ravishanker, GS.; Vijakumar, S.; Beveridge, DL. STRIPS: an algorithm for generating two-dimensional hydrogen bond topology diagrams for proteins. In: Smith, DA., editor. Modeling the Hydrogen Bond. American Chemical Society; Washington, DC: 1994. p. 209-219.

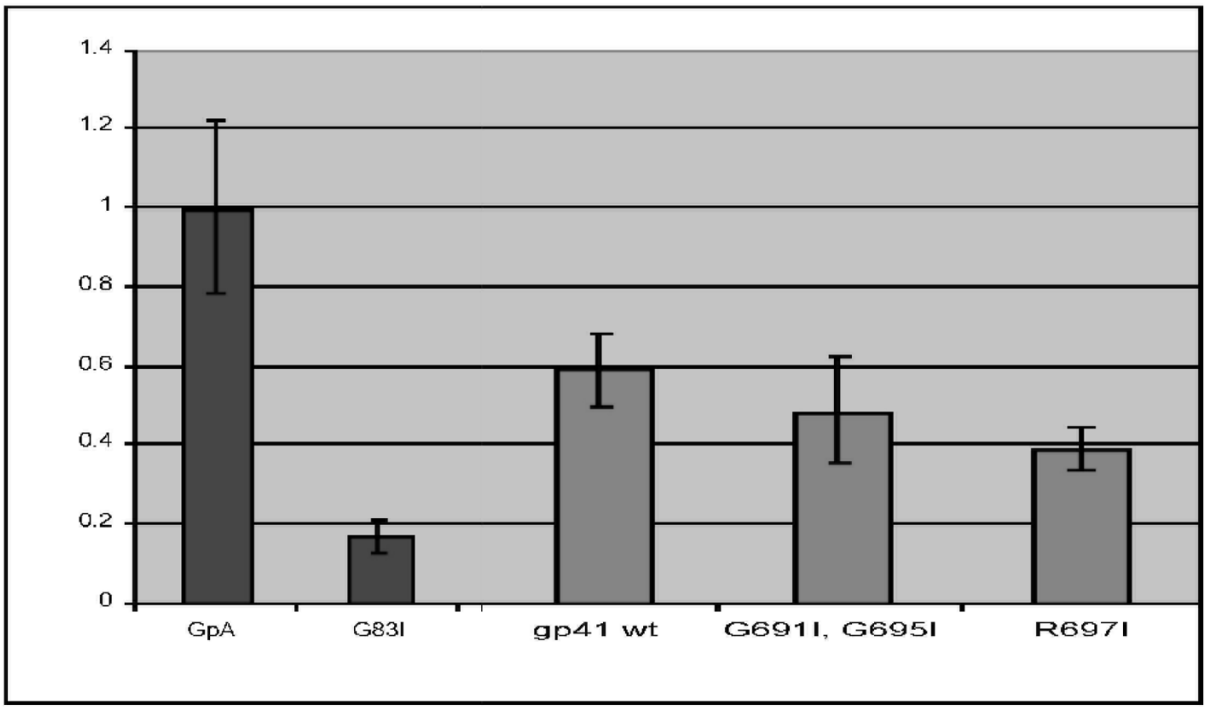
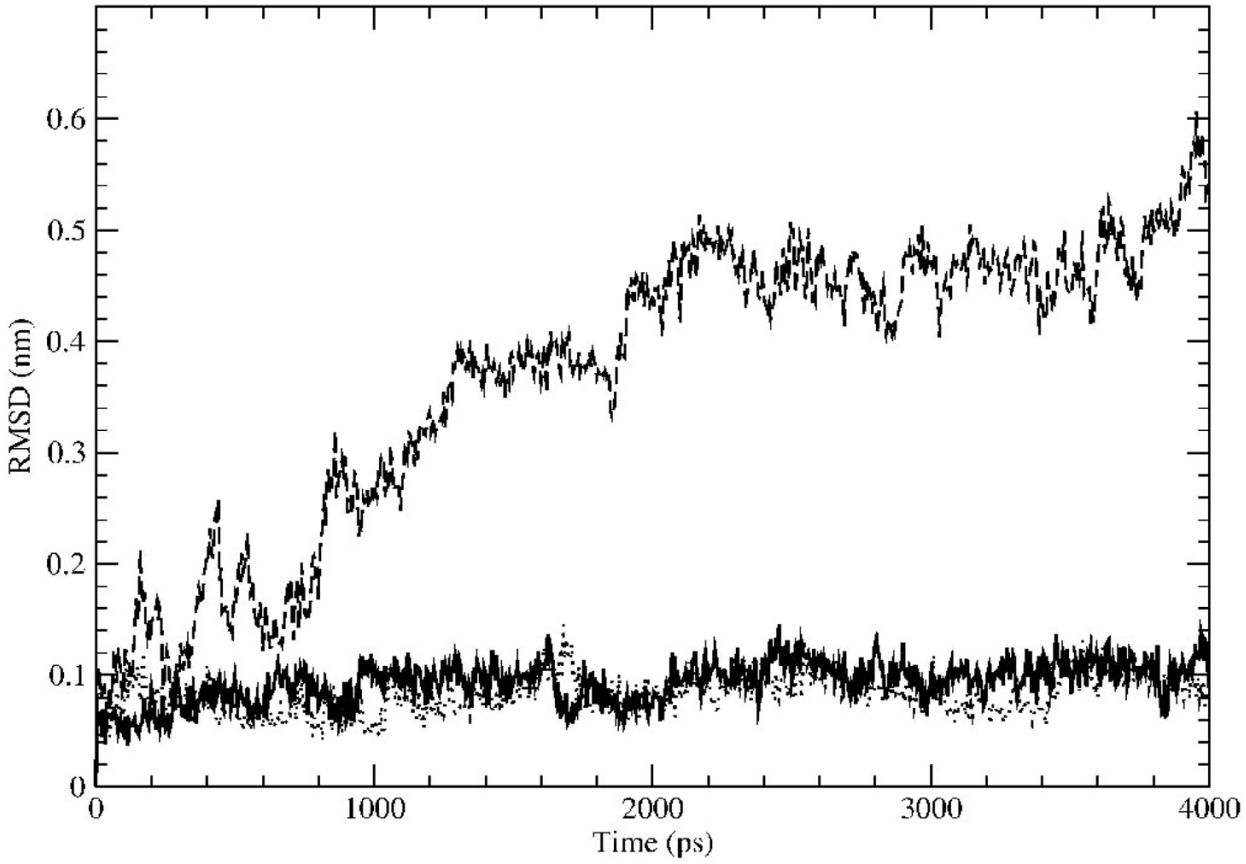
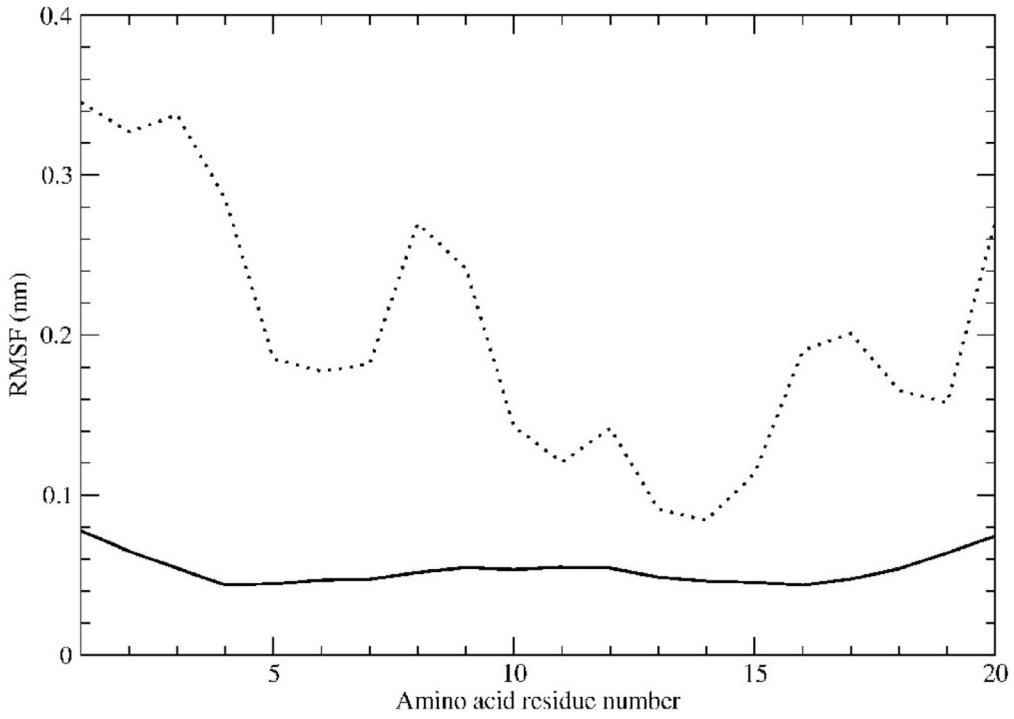
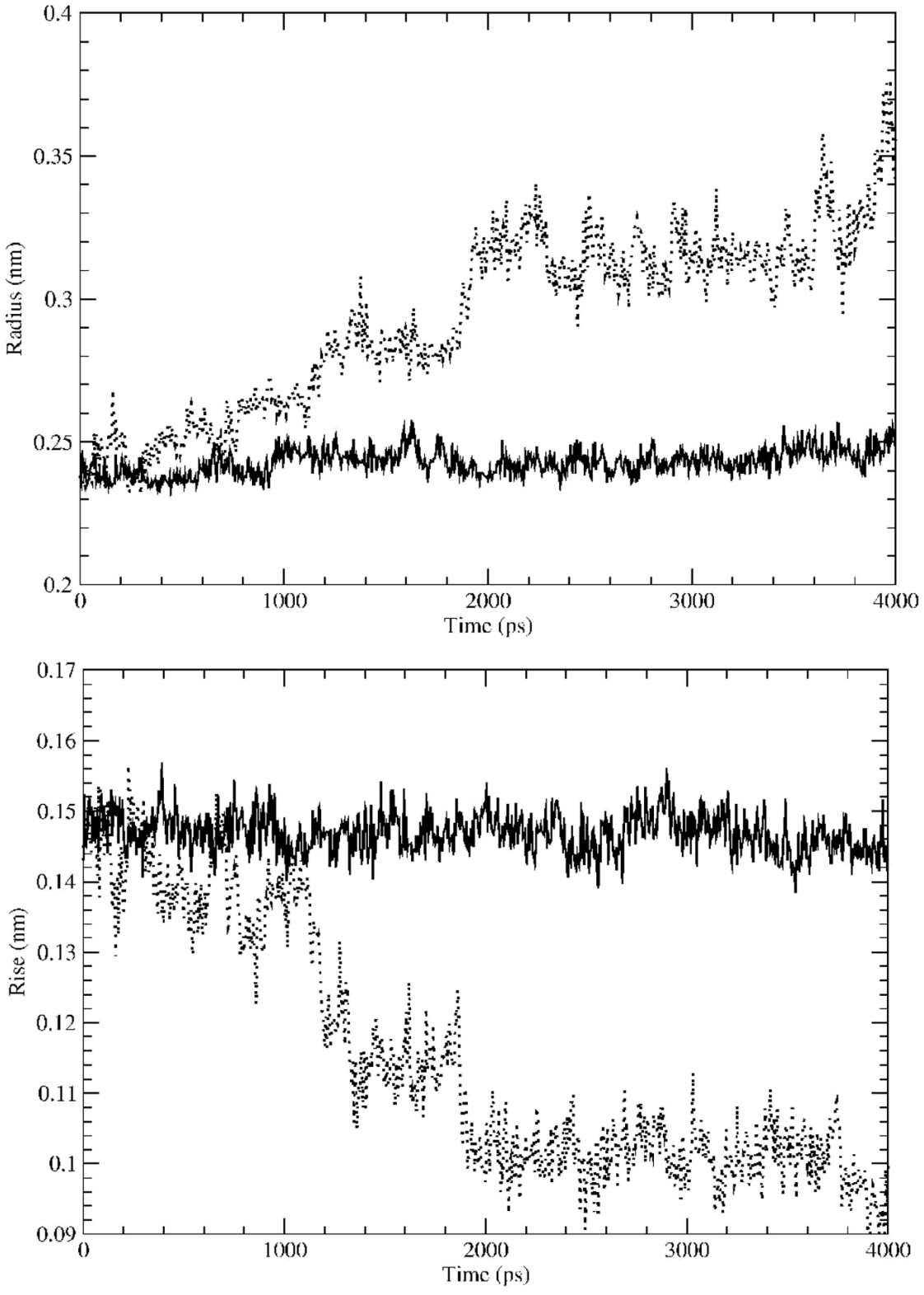


Fig. 1.

CAT levels normalized to the wild-type of glycophorin A (GpA). G83I is the dimer-disruptive mutant of glycophorin A. Gp41 wt shows significant CAT levels, implying the gp41 TM domains oligomerize. Upon changing the interfacial arginine residue to isoleucine residue, the CAT levels decrease.







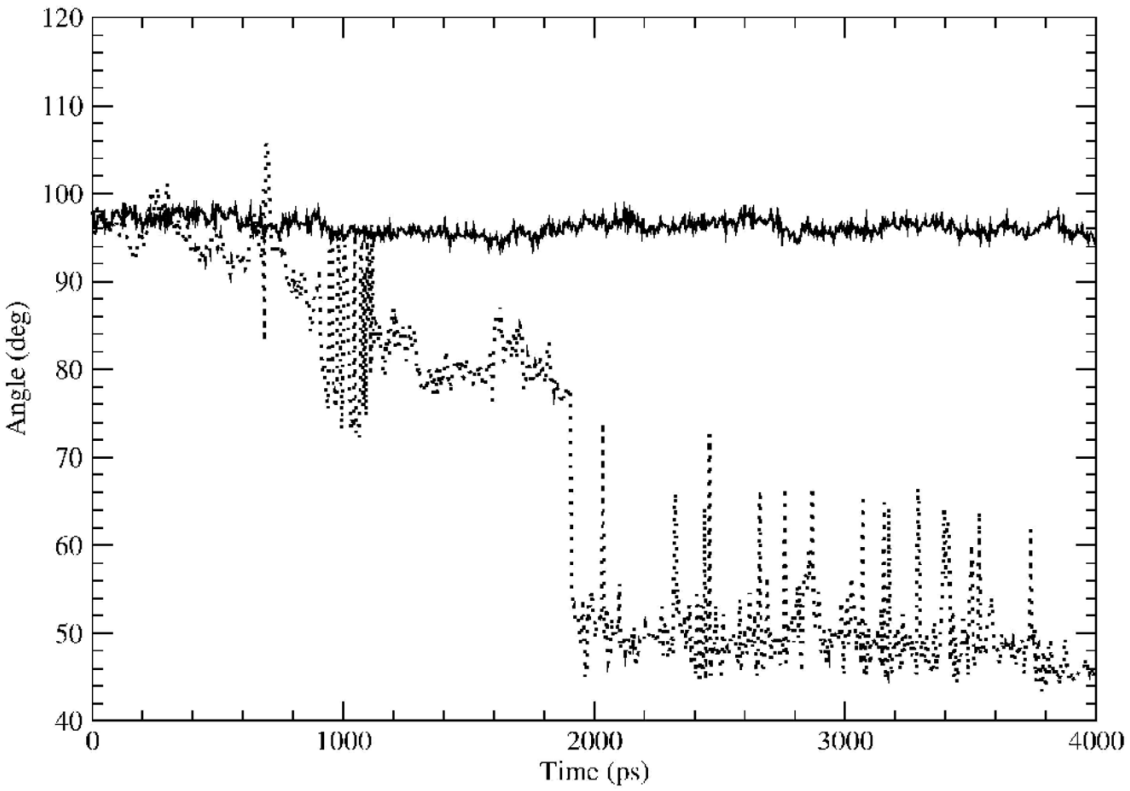
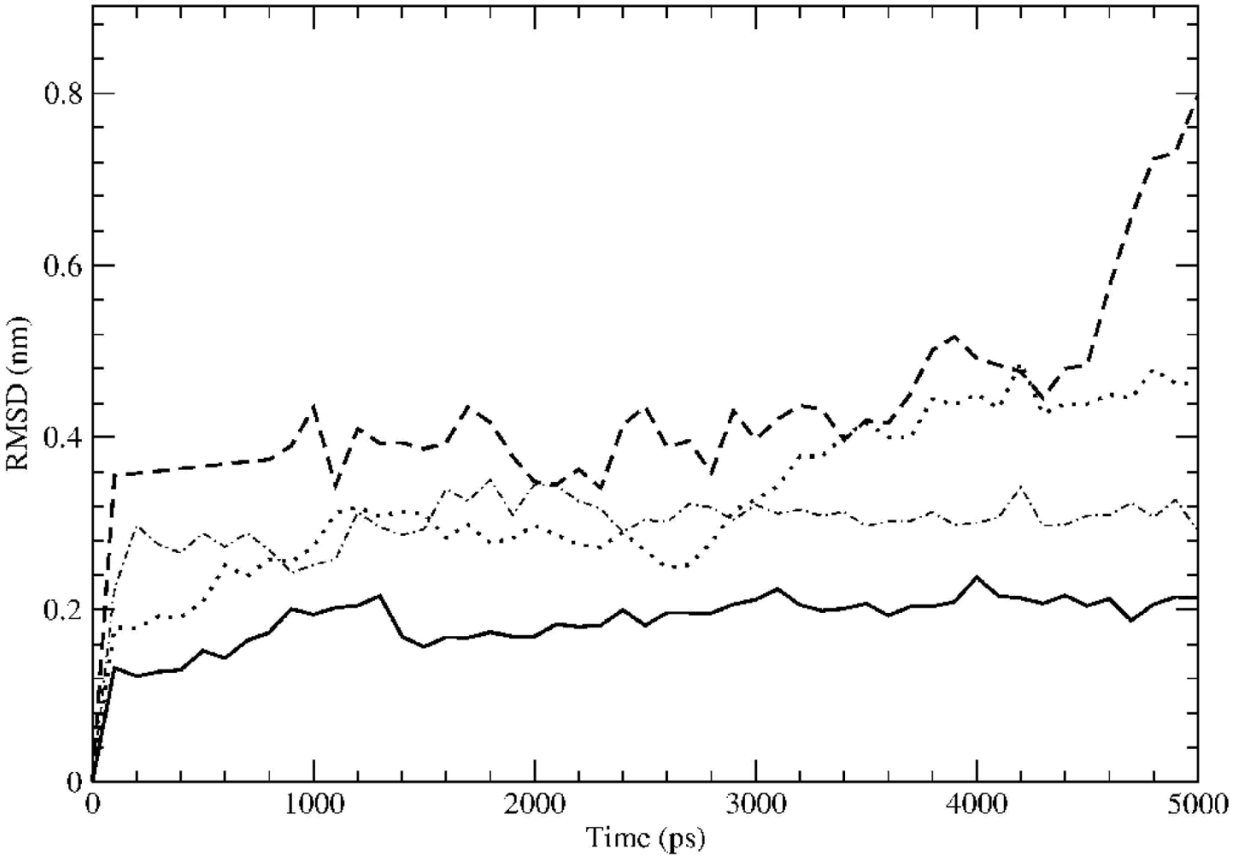
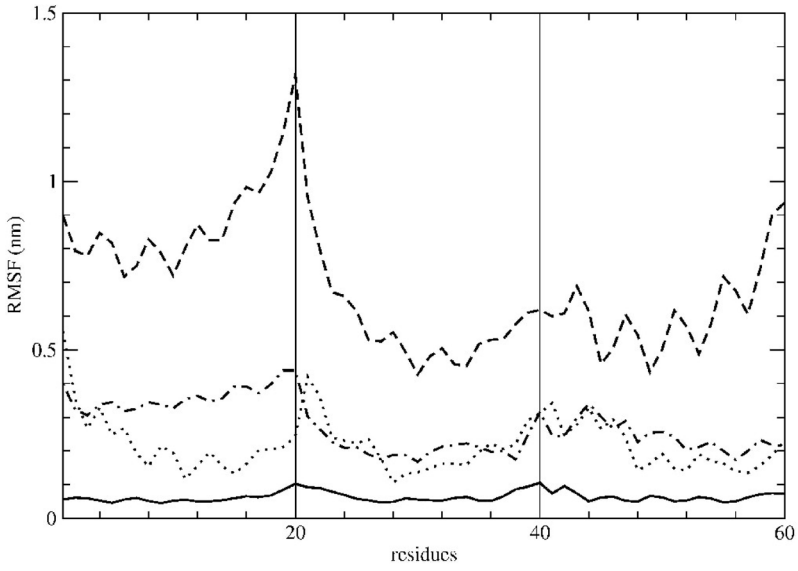


Fig. 2.

(a) C_α RMSD versus time for the 20-residue monomeric helix of gp41 TM domain in lipid (thick solid) and in water (dashed). C_α RMSD versus time for the 28-residue monomeric helix of gp41 TM domain in lipid (dotted) virtually overlaps with the 20-residue monomeric helix (thick solid). (b) C_α RMSFs for the same model in lipid (solid) and in water (dotted). Helix radius (c), rise per residue (d), and twist per residue (e) versus time for the 20-residue monomeric helix of gp41 TM domain in lipid (solid) and in water (dotted).





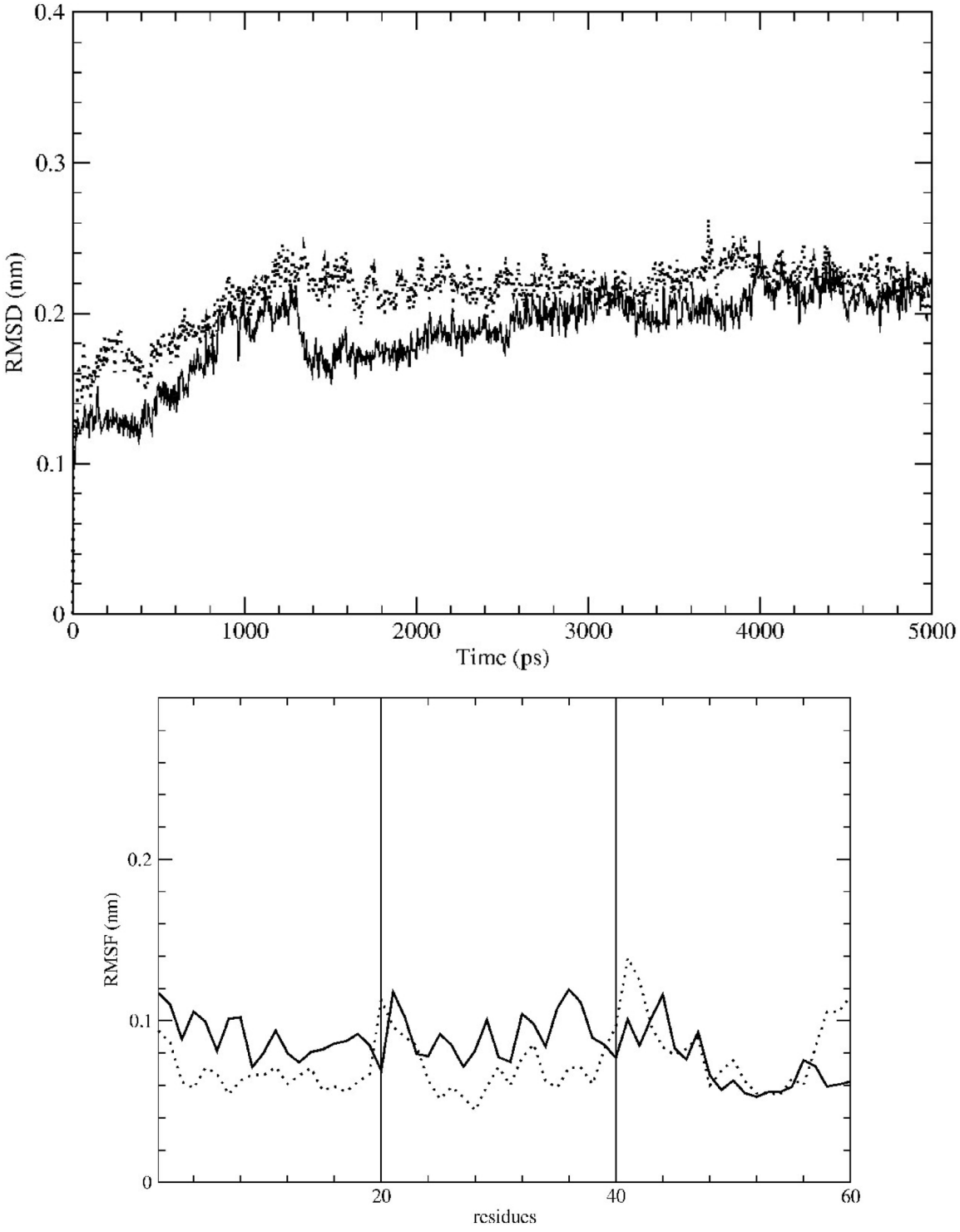


Fig. 3.

C_α RMSD versus time (a) and C_α RMSFs (b) for the three-helix bundle of gp41 TM domains in lipid (solid), in DMSO (dotted), in Decane (dashed), and in water (dot-dashed). C_α RMSD versus time (c) and C_α RMSFs (d) for right- (solid) and left-handed (dotted) three-helix bundle of gp41 TM domains in lipid.

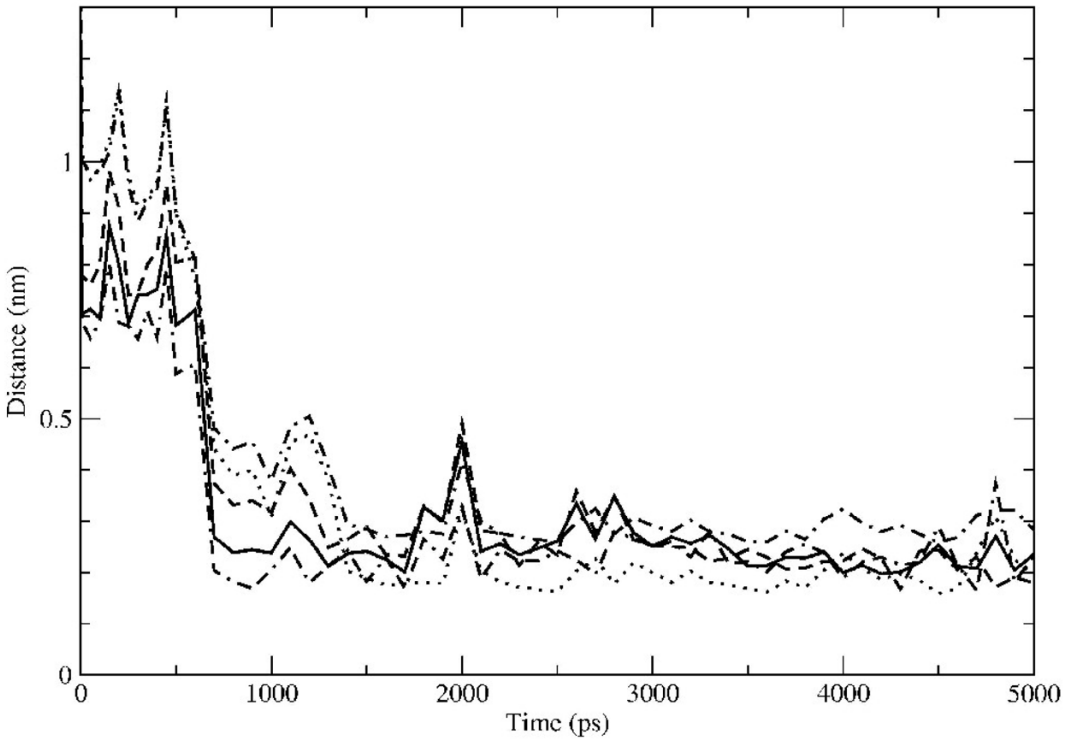
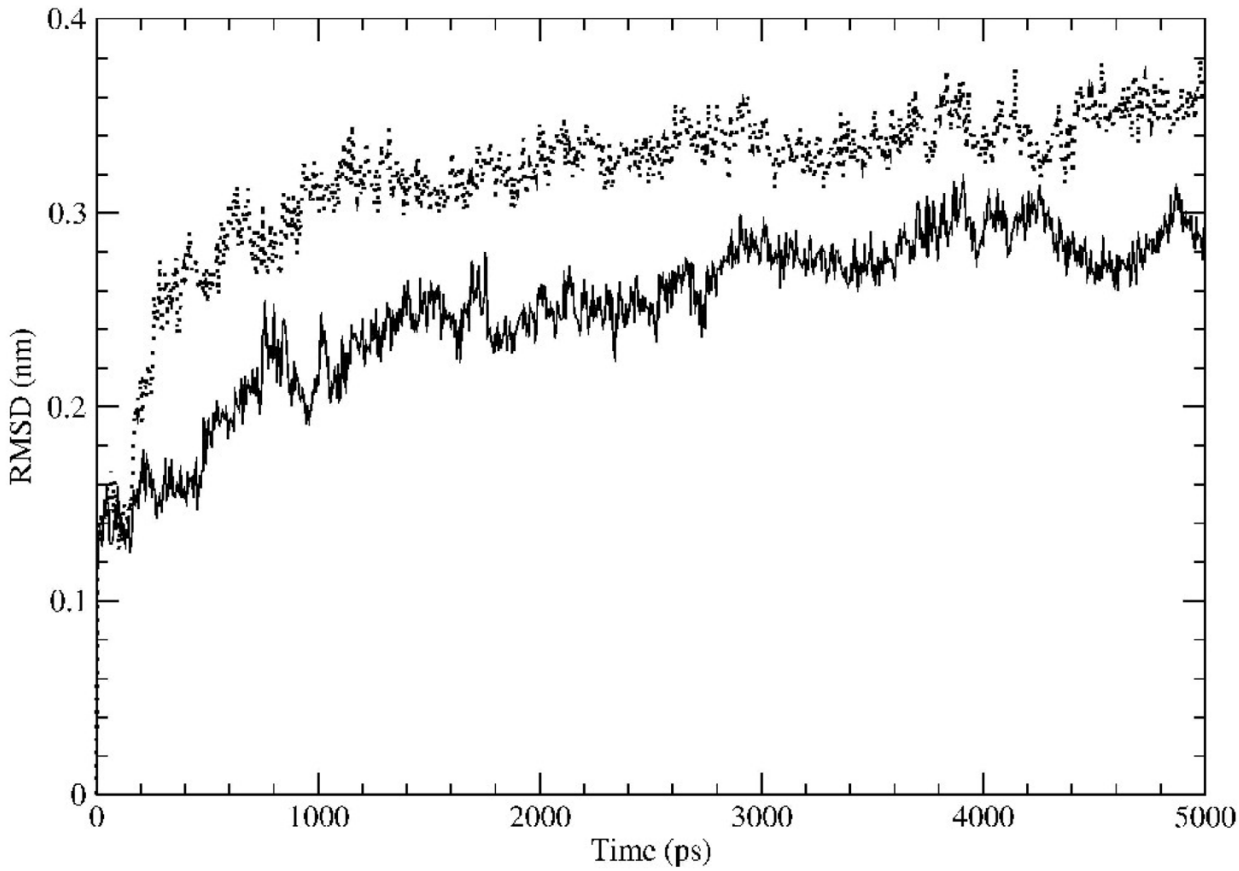
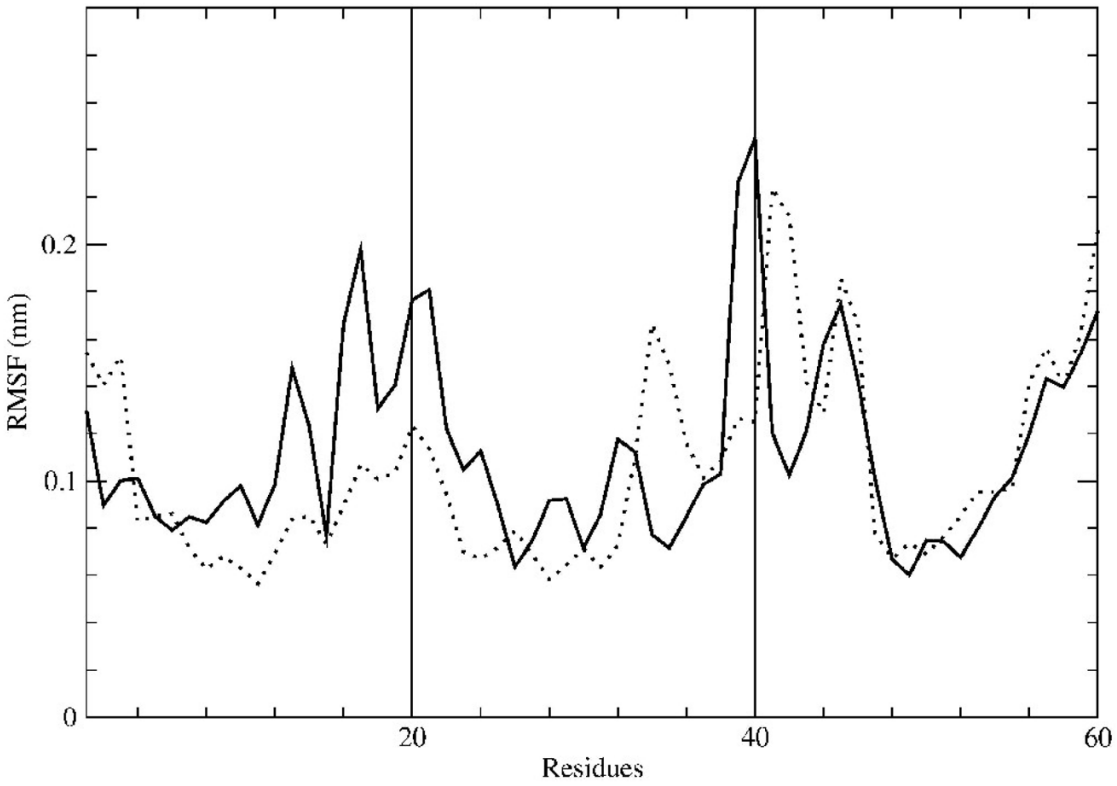
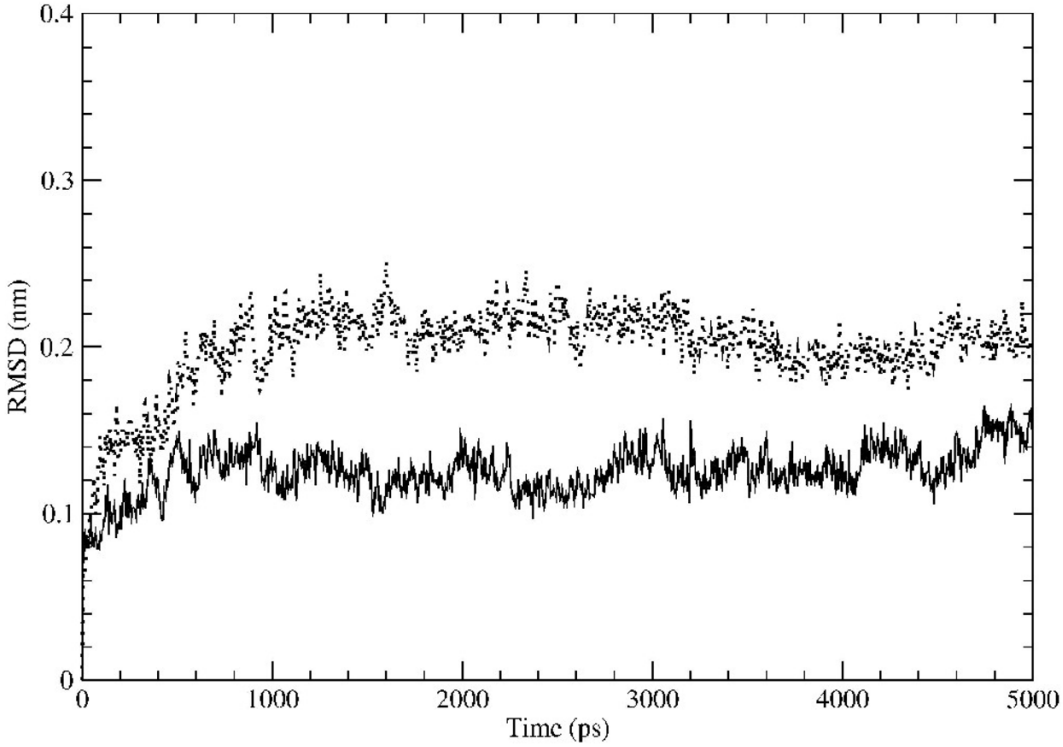


Fig. 4.

Several inter-chain bond distances during the first 5 ns of simulation of the three-helix bundle of gp41 TM domains in lipid. Different hydrogen bonds are shown: the hydrogen bond between Argn12 (C_ζ) of Helix I and the hydrogen of Argn12 (N_ε) of Helix III (solid); between the hydrogen of Argn12 (N_{H1}) of Helix I and Argn12 (N_{H1}) of Helix III (dotted); between Argn12 (N_{H1}) of Helix I and the hydrogen of Argn12 (N_ε) of Helix III (dashed); between Argn12 (N_{H1}) of Helix I and the hydrogen of Argn12 (N_{H1}) of Helix III (dot-dashed); and between Argn12 (N_{H2}) of Helix I and the hydrogen of Argn12 (N_ε) of Helix III (dot-dash-dashed).







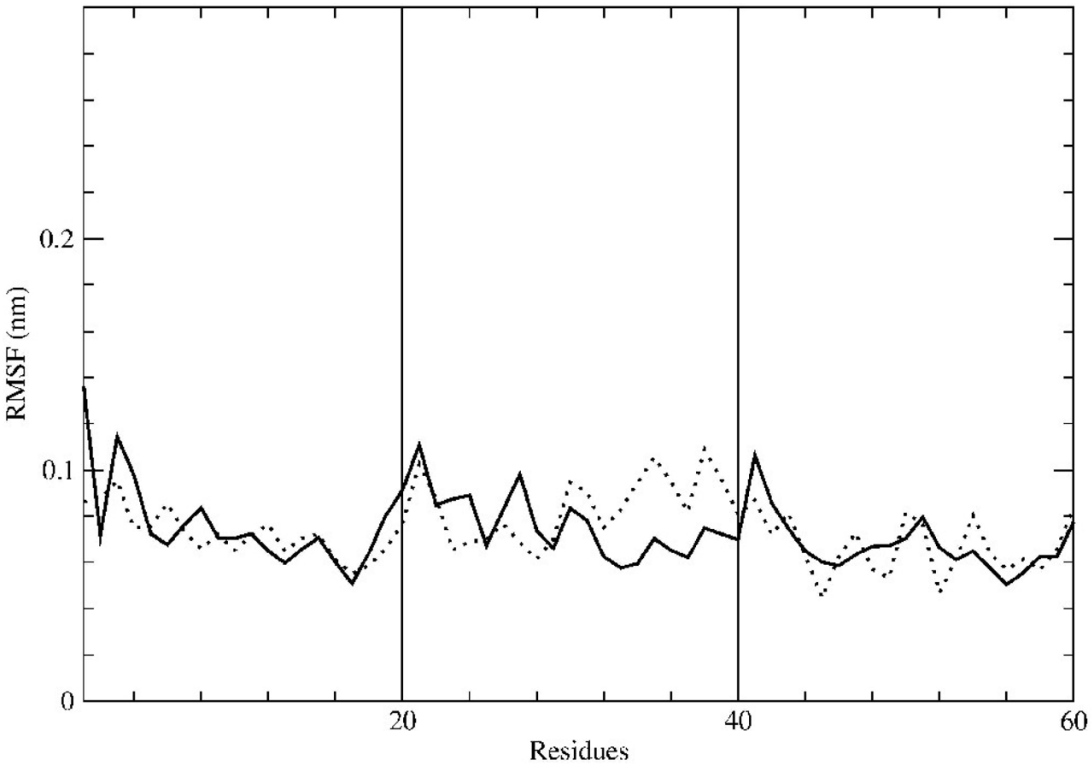


Fig. 5.

C_α RMSD versus time (a) and C_α RMSFs (b) for right- (solid) and left-handed (dotted) three-helix bundle of gp41 of wild-type with the protonated central arginine residues TM domains in lipid. C_α RMSD versus time (c) and C_α RMSFs (d) for right- (solid) and left-handed (dotted) three-helix bundle of gp41 of the R_I mutant TM domains in lipid.

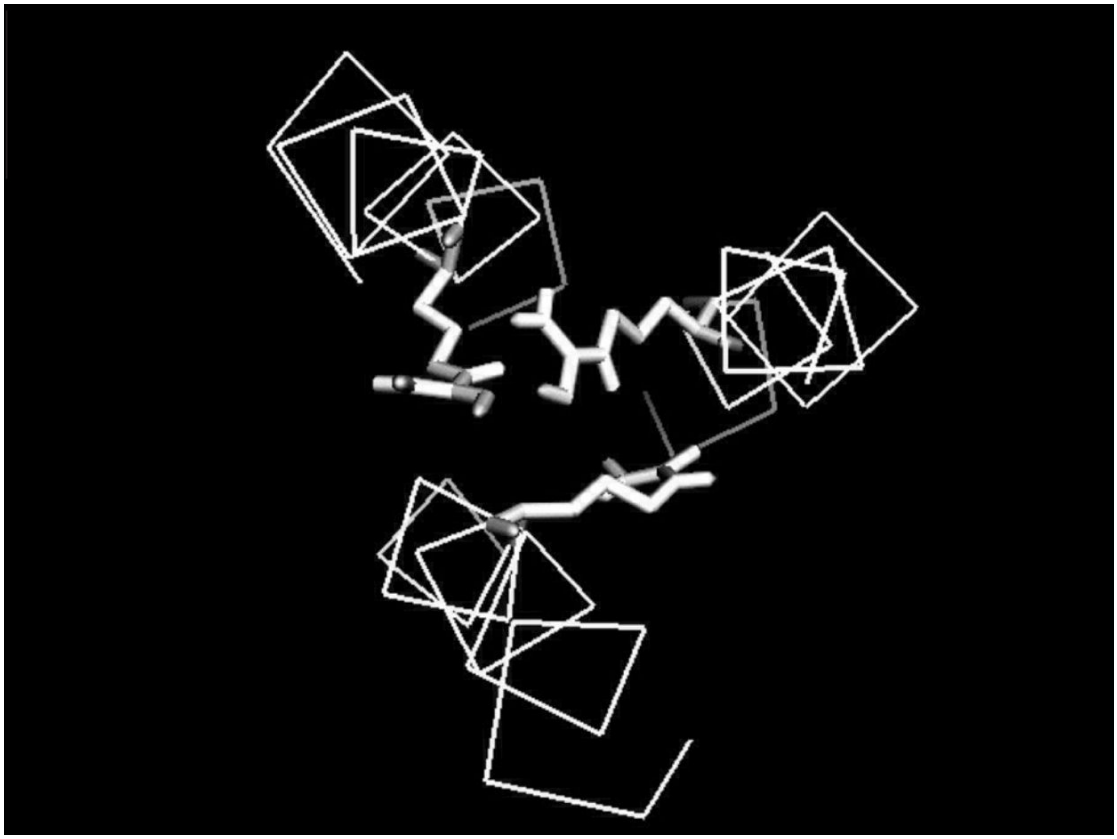


Fig. 6.
Snapshots of right-handed three-helix bundle viewed from top. Three deprotonated arginine (Argn) residues from each of the three helical chains are interacting with one another at the interfaces of the bundle.

Table 1

Stable inter-chain hydrogen bonds in the right-handed three-helix bundle of the 20-residue TM domain of GP41 from HIV-1 observed during the course of 13-ns simulations (after the 5-ns initial equilibration period)

Hygrogen bond	Distance (Å) 2nd 13 ns	Angle (°)2nd 13 ns	Distance (Å) 3rd 5ns	Angle (°)3rd 5ns
H bond 1	2.10 ± 0.61	150.8 ± 16.9	2.00 ± 0.35	150.0 ± 15.3
H bond 2	2.37 ± 0.69	131.9 ± 22.0	2.22 ± 0.31	137.2 ± 10.8
H bond 3	2.42 ± 0.56	141.9 ± 16.9	2.21 ± 0.24	151.2 ± 15.0
H bond 4	2.41 ± 0.49	118.2 ± 23.0	2.28 ± 0.35	123.4 ± 16.5
H bond 5	2.78 ± 0.39	127.2 ± 19.2	2.61 ± 0.31	142.0 ± 19.1

H bond 1: Argn12 (N_{H1}) of Helix I-H ····· Argn12 (N_{H1}) of Helix III

H bond 2: Argn12 (N_{H1}) of Helix I ····· H-Argn12 (N_{ε}) of Helix III

H bond 3: Argn12 (C_{ζ}) of Helix I ····· H-Argn12 (N_{ε}) of Helix III

H bond 4: Argn12 (N_{H2}) of Helix I ····· H-Argn12 (N_{ε}) of Helix III

H bond 5: Phe15 (C_{β}) of Helix I ····· H-Argn12 (N_{H1}) of Helix II

The bond distance is the X ··· H distance in angstroms (Å) shown as mean ± standard deviation, where X is the hydrogen bond acceptor. The bond angle is the X ··· H-Y angle in degrees shown as mean ± standard deviation, where X is the hydrogen bond acceptor and Y is the hydrogen bond donor. The criteria for defining a hydrogen bond are an average distance of X· ·· H less than 2.8 Å and an average angle of X ···H-Y greater than 120°, where X is the hydrogen bond acceptor [36]. The first 5 ns of the simulation is an equilibration phase, where the distances are very large at the beginning and gradually decrease to form the hydrogen bonds. Therefore, the mean distances have no meaning. Argn denotes deprotonated arginine.

Table 2

The lengths of the helices and the three-helix bundle, and the tilt angles of the helices and the bundle normal to the lipid bilayer, in the course of MD simulations

Helix Simulation	**wild type (RH) 23 ns**		**wild type (LH) 18 ns**		**R_I (RH) 15 ns**	
	length (nm)	**tilt (°)**	**length (nm)**	**tilt (°)**	**length (nm)**	**tilt (°)**
Chain 1	2.87 ± 0.04	10.7 ± 4.5	2.90 ± 0.05	19.5 ± 2.9	2.93 ± 0.10	16.9 ± 3.1
Chain 2	2.83 ± 0.06	12.3 ± 2.7	2.99 ± 0.06	18.4 ± 3.6	2.80 ± 0.05	12.1 ± 3.2
Chain 3	2.71 ± 0.05	15.3 ± 4.8	2.87 ± 0.07	30.0 ± 3.7	2.79 ± 0.05	7.6 ± 3.7
Bundle	2.77 ± 0.04	9.9 ± 3.4	2.71 ± 0.04	7.0 ± 3.2	2.79 ± 0.03	3.9 ± 2.4
Helix Simulation	R_I (LH) 15 ns		wild type, arg12 (RH) 13 ns		wild type, arg12 (LH) 10 ns	
	length (nm)	tilt (°)	length (nm)	tilt (°)	length (nm)	tilt (°)
Chain 1	3.01 ± 0.08	20.7 ± 2.6	2.60 ± 0.11	25.3 ± 4.3	2.92 ± 0.09	17.6 ± 6.9
Chain 2	2.85 ± 0.06	9.6 ± 3.0	2.82 ± 0.06	21.9 ± 5.6	3.02 ± 0.07	18.5 ± 3.4
Chain 3	2.88 ± 0.08	17.1 ± 3.3	2.81 ± 0.10	17.2 ± 3.9	2.71 ± 0.08	45.9 ± 5.5
Bundle	2.80 ± 0.04	5.2 ± 2.6	2.56 ± 0.06	6.4 ± 2.9	2.54 ± 0.08	9.8 ± 2.6

The values are shown as mean ± standard deviation.