Molecular Dynamics Simulations of the E1/E2 Transmembrane Domain of the Semliki Forest Virus

Ana Caballero-Herrera and Lennart Nilsson
Department of Biosciences at Novum, Karolinska Institutet, S-141 57 Huddinge, Sweden

ABSTRACT Transmembrane (TM) helix-helix interactions are important for virus budding and fusion. We have developed a simulation strategy that reveals the main features of the helical packing between the TM domains of the two glycoproteins E1 and E2 of the α-virus Semliki Forest virus and that can be extrapolated to sketch TM helical packing in other α-viruses. Molecular dynamics simulations were performed in wild-type and mutant peptides, both isolated and forming E1/E2 complexes. The simulations revealed that the isolated wild-type E1 peptide formed a more flexible helix than the rest of peptides and that the wild-type E1/E2 complex consists of two helices that intimately pack their N-terminals. The residues located at the interhelical interface displayed the typical motif of the left-handed coiled-coils. These were small and medium residues as Gly, Ala, Ser, and Leu, which also had the possibility to form interhelical C_α-H$\cdots$O hydrogen bonds. Results from the mutant complexes suggested that correct packing is a compromise between these residues at both E1 and E2 interhelical interfaces. This compromise allowed prediction of E1-E2 contact residues in the TM spanning domain of other alphaviruses even though the sequence identity of E2 peptides is low in this domain.

INTRODUCTION

Alphaviruses form by budding from the plasma membrane of an infected cell and enters uninfected cells by a membrane fusion event. Although the mechanism of these events is still unclear, recent studies have pointed out that they are controlled by the viral membrane protein heterodimers E1/E2 (Garoff and Cheng, 2001). It has been proposed that the transmembrane (TM) segments of the spikes, which are arranged as pentamers and hexamers in the viral envelope at neutral pH, reorganize to form trimers at the low pH characteristic of virus entry. It is this reorganization that triggers the membrane fusion process (Haag et al., 2002). To understand the fusion mechanism, it is necessary to know about the interactions of the heterodimer E1/E2 in the TM domain. The study of TM protein interactions has however been a challenging task during the last decades since high resolution structural data are difficult to obtain for non-soluble membrane proteins. For this reason, the insight that computational methods provide is of great importance. Molecular dynamics (MD) simulations is probably the most widely used tool to analyze TM peptides and proteins at the atomic level. Of particular interest is the information about the lipid/protein interaction supplied by simulations with explicit lipid bilayer (Duneau et al., 1999; Forrest et al., 2000; Law et al., 2000; Petrache et al., 2000; Woolf, 1997, 1998) or membrane mimetics (Bright and Sansom, 2003). Furthermore, different methods have been developed to predict the topology and packing of TM proteins (Adams et al., 1996; Krogh et al., 2001; Pappu et al., 1999).

As a result of these computational studies, together with experimental work, a wide knowledge about TM proteins is available. It is known that the membrane spanning domains (which are encoded by ~30% of most genomes) are typically formed by hydrophobic helical segments of ~20–24 amino acids, and their folding mechanism has been postulated in the two-stage model (for review see Popot and Engelman, 2000). In this model, the partially formed helices are first inserted and stabilized individually inside the lipid bilayer and only then they associate to form helix bundles (Henderson, 1975, 1977; Popot and Engelman, 1990; Singer, 1990; Singer and Yaffe, 1990). In the absence of interactions with water, other interactions such as the intra and interhelical hydrogen bonds, ion pairs, dipole-dipole interactions between helices, and the lipid/protein interactions have to overcome the unfavorable loss in entropy of keeping the TM helices together. Also the van der Waals interactions become very important here. In particular, these interactions are thought to be the promoter of the detailed close packing of TM α-helices (Popot and Engelman, 1990). It has been suggested that the amino acid sequence dictates the specific interactions that generate oligomerization (Laage et al., 2000; Lemmon et al., 1994; Sternberg and Gullick, 1989, 1990; Whitley et al., 1993), and different sequence motifs that may enhance homo- and heterodimerization have been proposed. Residues such as Gly, Ala, Ser, Thr, Leu, Val, Ile, and Met have a fundamental importance (Eilers et al., 2000; Liu and Deber, 1998; Sternberg and Gullick, 1990), in apparent contradiction to the low helical propensity that some of these amino acids show in aqueous solution. In particular, the structural role that glycine seems to play in TM α-helices is remarkable. The role as helix breaker in soluble proteins of glycine contrasts with its role as molecular notch for helix-helix packing at the interhelical interfaces in polytopic membrane proteins (Javadpour et al., 1999). It has even been shown that the energy of the Gly-83-Gly-83

Submitted February 14, 2003, and accepted for publication September 3, 2003.

Address reprint requests to Lennart Nilsson, E-mail: lennart.nilsson@biosci.ki.se.

0006-3495/03/12/3646/13 $2.00

interaction, which is thought to be crucial for the glycophorin A (GpA) dimerization, is electrostatically unfavorable (Petrache et al., 2000).

Due to the low dielectric constant inside the apolar lipid bilayer, the intra- and interhydrogen bonds become so strong that the entropic penalty of restraining in a helical configuration residues with very low helical propensity (Popot and Engelman, 2000) and other unfavorable interactions can become overcome. Moreover, even the presumably weak C_α-H$\cdots$O hydrogen bond gains importance inside the low dielectric lipid bilayer environment. Its energy in vacuo lies between 2.5 and 3.0 kcal/mol, approximately half of that of a common amide hydrogen bond (Scheiner et al., 2001). Thus, it is logical to think that inside the membrane the commonly found networks of C_α-H$\cdots$O hydrogen bonds at the interhelical interfaces may be critical factors for helical association stability as well as for specificity (Fleming and Engelman, 2001; Senes et al., 2001).

Here we investigate by molecular dynamics simulations the interactions between the two membrane-spanning domains of the glycoproteins E1 and E2 of the alphavirus Semliki Forest virus (SFV), which experimentally are known to form helices packed together in a left-handed fashion (Mancini et al., 2000). The TM segments of the wild-type glycoproteins E1 and E2 individually and in complex as well as other E1/E2 complexes where some residues, mainly glycines, of E1 or E2 or both were mutated to leucines have been studied (Tables 1 and 2). Some of these complexes have been investigated experimentally and show defects in both heterodimer stability and virus budding (envelope assembly) (Sjöberg and Garoff, 2003).

We have developed a simulation strategy that characterizes the structural differences between the mutant and wild-type heterodimers, as well as it identifies the amino acids located at the helical interface that may promote the E1/E2 heterodimerization. Our results correctly predict defects in heterodimer interactions that are manifested as decreased heterodimer stability in the virus (Sjöberg and Garoff, 2003). We believe that our method can discriminate between correct and improper packing of the TM domain of this alphavirus, and that it can be used as a first step in the design of new mutants as well as to roughly sketch the TM spanning domain interactions of the E1/E2 complex of other alphaviruses.

METHODS

MD simulation protocol

The CHARMM (Brooks et al., 1983) program with the all-atom parameter set (MacKerell et al., 1998) was used in all simulations. An atom-based force-shift method for the long-range electrostatic interactions, which is known to produce accurate and stable simulations (Norberg and Nilsson, 2000), and an atom-based shifting function for the van der Waals interactions were used to truncate the nonbonded interactions at 12 Å. The nonbonded list was generated with a cutoff of 13 Å and updated as soon as any atom had moved 0.5 Å or more. Since the total number of atoms in the simulations was very small, the use of a longer cutoff for the nonbonded interactions could be considered. Thirty angstroms is the standard lipid bilayer thickness and two 1.5-ns simulations with cutoffs 18 Å and 30 Å of the $E1^{WT}/E1^{WT}$ complex were performed. The 18-Å and 30-Å cutoffs increased the CPU time compared to the 12-Å cutoff by 2.7 and 3.6, respectively, and the results did not vary essentially from those with 12-Å cutoff. The backbone root mean square deviation (RMSD) evolution of the three trajectories was similar with higher fluctuations for the largest cutoff case, but it remained under 1 Å most of the time (data not shown). Also the residues located at interhelical interfaces coincided in the three cases.

SHAKE was applied to all covalent bonds involving hydrogens (Ryckaert et al., 1977). Vacuum ($\varepsilon = 1$) was chosen as dielectric medium to mimic the lipid bilayer. MD simulations with explicit lipid bilayers have shown that the structure, dynamics, and energetics of individual α-helices as well as α-helical dimers depend on the lipid bilayer (Petrache et al., 2000; Woolf, 1997, 1998); however, the results of the GpA dimer and four different explicit lipid bilayers from Woolfs group suggest that the dimer average structure does not change significantly with the lipid (Petrache et al., 2000). In our simulations the lipid bilayer is substituted by vacuum because of the obvious reduction of the simulation time and because an average structure is enough for our purposes, but one should be aware of the simplification introduced here. Principally, as pointed out by Petrache and co-workers (Petrache et al, 2000), this approximation will neglect the lipid modulation of the protein fluctuation about the average structure, that is, the specific flexibility of the protein.

Before the simulation, all the starting configurations were subjected to a gradual minimization consisting of 50 steepest-descent (SD) steps of minimization with harmonic constraints on all atoms with a force constant of 20 kcal/mol/Å^2, followed by 50 steps of an adopted basis Newton-Raphson (ABNR) minimization . At this point, the harmonic force constant was reduced to half and only applied to the backbone in the next 50 steps of ABNR minimization. Finally the harmonic constrains were completely turned off during the last 1100 steps of ABNR minimization. The leapfrog algorithm was used in all simulations. Each simulation was initialized with a 5-ps heating period where the velocities were increased in increments of 5 K every 0.1 ps followed by a 5-ps equilibration period. The temperature in the equilibration period was checked every 0.1 ps and constrained to be 300 $\pm$ 10 K by scaling the velocity. The integration time step was 2 fs and the coordinates were saved every 0.1 ps for analysis. Production trajectories were performed without velocity scaling during 20.1 ns, in the single-helix case, and 2.1 ns for the helix-helix complexes. It may be argued that the complex simulations are too short to insure convergence. However, longer trajectories (up to 20.1 ns) were produced for some of the helix-helix complexes. The results were similar to those obtained from the 2.1

TABLE 1 Amino acid sequences of the individual putative TM helices

E1	WT	ISGGLGAFAIGAILVLVVVTCIGL
	4L	ISLLLLAFAILAILVLVVVTCIGL
	11L	LLLLLLLLLLLLILVLVVVTCIGL
E2	WT	SAVVGMSLLALISIFASCYMLVAA
	AN	SAAVGMSLLALISIFASCYMLVAA
	3L	SLAVLMSLLLLISIFASCYMLVAA

TABLE 2 E1/E2 complexes

	E2		
E1	WT	AN	3L
WT	$E1^{WT}/E2^{WT}$	$E1^{WT}/E2^{AN}$	$E1^{WT}/E2^{3L}$
4L	$E1^{4L}/E2^{WT}$	$E1^{4L}/E2^{AN}$	$E1^{4L}/E2^{3L}$
11L	$E1^{11L}/E2^{WT}$	$E1^{11L}/E2^{AN}$	

simulations (the residues located at the interhelical interfaced are shown in Fig. 4 as *open triangles* and coincide with those obtained from the shorter simulations).

Sequences

The Cryo-Electron Microscopy study of Mancini et al. (2000) provides experimental structure data of the SFV TM domain at low resolution. It suggests that the TM domain of the virus consists of two α-helices packed together in a parallel left-handed fashion (see Fig. 1).

However this study neither provides coordinates nor informs accurately about which amino acids of glycoproteins E1 and E2 constitute the putative TM α-helices. In our study, segments of 24 amino acids, with acetylated N-terminal and amidated C- terminal, were considered to be completely immersed inside the membrane (Table 1). This decision was supported by the results obtained with the membrane protein topology predictor TMHMM (Krogh et al., 2001). Eight complexes composed of wild-type or modified E1 and E2 peptides (Tables 1 and 2) were studied.

The E1 variants were two peptides with 4 (4L) or 11 (11L) leucines introduced in the N-terminal part of E1, which is close to the external surface of the membrane in the virus. Experimental data about viral particle formation and heterodimer stability in the virus of the wild-type and the mutant complexes $E1^{4L}/E2^{WT}$ and $E1^{11L}/E2^{WT}$ is available (Sjöberg and Garoff, 2003) but not for the complexes containing the E2 variants AN and 3L. These E2 mutants had one Val-Ala substitution and the 3L three more residues mutated to leucines.

Initial coordinates

Three different procedures were used to assign initial coordinates: 1), The single helix initial coordinates were built in the program CHARMM as an ideal α-helix. 2), In the case of the putative helix-helix complexes the determination of the first starting coordinates follows a more sophisticated process. Twenty starting configurations were generated for each of the complexes $E1^{WT}/E2^{AN}$, $E1^{4L}/E2^{AN}$, $E1^{11L}/E2^{AN}$ named *state i* ($i = 1,20$). The first set of coordinates of each complex was obtained from the glycophorin A TM helix dimer (pdb code 1AFO) (MacKenzie et al., 1997; Treutlein et al., 1992), whose helices interact in a parallel right-handed fashion. The other 19 states were derived from this first one by a screwlike motion of the helix E1 onto the helix E2; the missing side-chain coordinates were built in CHARMM in an extended conformation.

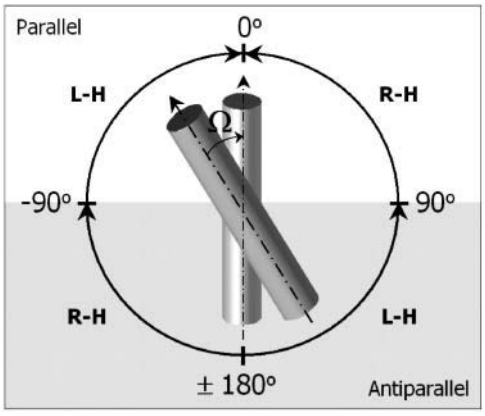

FIGURE 1 Helix-helix packing criteria. Two parallel helices pack in a right-handed (*R-H*) fashion if the interhelical angle Ω is between 0° and 90° or between 180° and 270° and in a left-handed (*L-H*) fashion for Ω angles between 0° and −90° or between 90° and 180°.

The third coordinate generation procedure was used for the $E1^{WT}/E2^{WT}$, $E1^{4L}/E2^{WT}$, $E1^{11L}/E2^{WT}$, $E1^{WT}/E2^{3L}$, and $E1^{4L}/E2^{3L}$ complexes. Ten states were generated for each complex. The 10 backbone starting coordinates of each complex were the final structures of 10 of the 20 states of the $E1^{WT}/E2^{AN}$ complex chosen after the analysis of their trajectories. The remaining 10 states were neglected as starting coordinates for two reasons a), because the helices of this states displayed very poor contacts and b), because their packing would be impossible due to the limitations imposed by the lipid bilayer. The helices that do not pack in a very parallel manner were neglected; otherwise the helices would not be totally embedded in the bilayer since a typical lipid bilayer is ~30-Å thick.

Analysis procedures

The backbone root mean square deviation (RMSD) matrix is built up calculating the backbone RMSD between every conformation and all the other conformations at 20-ps intervals.

The elements of the root mean square fluctuation (RMSF) matrices were constructed for all pairs of C_α atoms by calculating the RMSF of the interatomic distances between the C_α atoms sampled at 20-ps intervals during the simulations.

The C_α contact map of a specific E1/E2 complex is a 24 × 24 matrix where each *ij* cell contains the mean distance during the last 1.1 ns of simulation between the $C_{\alpha i}$ atom of E1 and the $C_{\alpha j}$ atom of E2. Two C_α atoms are considered to form a contact if their distance is less than 7 Å. The C_α probability contact matrix was constructed from the C_α contact maps of the chosen simulations of a particular E1/E2 complex. It is also a 24 × 24 matrix where each a_{ij} value is the normalized number of times that the E1 $C_{\alpha i}$ atom makes a contact with the E2 $C_{\alpha j}$ atom among the simulations taken into account for that complex. The simulations that lead to an antiparallel orientation between helices or those that possessed unfeasible contacts were discarded.

In the helix packing curves the minimal distances between interhelical C_α atoms are displayed only if they are lower than 6 Å.

The occurrence of C_α-H···O hydrogen bonds was calculated over the last 1.1 ns of simulation of every simulation. From these was obtained the average occurrence of every particular C_α-H···O hydrogen bond for a particular complex over all simulations of that complex.

RESULTS AND DISCUSSION

Single helix dynamics

Our first approach to determine the structural differences between the different viruses consisted of the simulation of the isolated TM segments of the glycoproteins E1. As predicted by the two-stage model (Popot and Engelman, 1990), the segments displayed helical structure. All the segments showed a significant high π-helical content after the first 1 ns through the whole simulation. Similar behavior is exhibited in other MD studies in vacuum and in solvated lipid bilayer environment (Duneau et al., 1999, 1996; Lee et al., 2000). π-Helical segments, sometimes called α-aneurisms, have been found experimentally (Keefe et al., 1993; Morgan et al., 2001; Rajashankar and Ramakumar, 1996), and they are believed to be intimately involved in protein function (Weaver, 2000). In the case of TM helices, π-helical turns seem necessary to accommodate long residue strings inside the lipid bilayer (Popot and Engelman, 2000). However, a recent work (Feig et al., 2003) suggests that the formation of π-helical segments in MD is a force field

artifact. In fact they show that different force fields bias toward π-helical formation when solvation is taken into account, whereas α- and π-helices are energetically equivalent in vacuum. The O(i)–HN(i + 5) main-chain hydrogen bonds normally survived 90% of the simulation time coexisting sporadically with conventional α-hydrogen bonds, but also with 3_{10}-hydrogen bonds, O(i)–HN(i + 3), at the end of the helices. The mutations introduced in the N-terminal of the helix uniquely had a local influence on the first third of the helix, avoiding the propagation of the disruption toward the C-terminal. Occasionally residues at the middle of the helices lost their main-chain hydrogen bonds, probably because of the formation of a kink, even though the lipid bilayer is treated as a continuum dielectric and not explicitly. The RMSD matrix (data not shown) of the E1 wild type and peptides 4L and 11L revealed that once the ideal α-helical structure imposed initially was lost the systems evolved gradually without any important structural change. Every conformation after the first nanosecond was correlated with the previous ones, with RMSD values below 2.6 Å. The calculated distance between the C_α atoms of residues 4 and 20 was approximately the same in all cases. The average distance corresponded to a helical rise per residue of 1.2 Å, which is closer to the 1.1 Å of a π-helix than to the 1.45 Å of a typical α-helix (Walther et al., 1996).

For most peptides the energy reached a stable minimum after 3 ns of simulation, the wild-type E1 peptide being the one with lowest energy (510 kcal/mol), followed by the 11L peptide (522 kcal/mol), then by the 4L peptide (546 kcal/mol). Finally we also calculated the RMSF matrix of the C_α atoms (Fig. 2). Although the fluctuations at the C_α atoms of the E1 wild-type peptide and the 4L peptide only correlated with the closest C_α atoms and in some degree with those

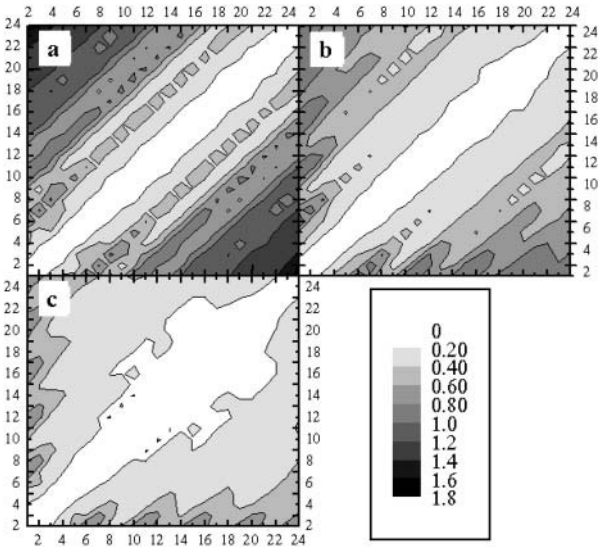

FIGURE 2 C_α RMSF matrices. (*a*) Wild-type E1 peptide, (*b*) 4L peptide, and (*c*) 11L peptide. The axes are labeled with the residue numbers.

eight amino acids apart, reflecting the α/π-helical character of the peptide, the rest of peptides displayed larger correlation regions. This finding supports the idea that the peptide 11L that contains only leucines in the first half of the helix confers a considerable rigidity to the helix. Previous MD studies (Bright and Sansom, 2003) also observe a slight increment of flexibility when the GxxG motif is introduced in a host polyalanine TM α-helix.

The E2 wild type was also simulated. No significant differences were found with respect to the E1 wild type.

E1-E2 helical complex dynamics

Conformational sampling

To study the interactions responsible for the helical association between the two putative helices E1 and E2, we performed molecular dynamic simulations on the E1/E2 complexes: $E1^{WT}/E2^{WT}$, $E1^{WT}/E2^{AN}$, $E1^{WT}/E2^{3L}$, $E1^{4L}/E2^{WT}$, $E1^{4L}/E2^{AN}$, $E1^{4L}/E2^{3L}$, $E1^{11L}/E2^{WT}$, and $E1^{11L}/E2^{AN}$ (see Table 2). Independent simulations of 2.1 ns were carried out for each complex, 10 in the case of the complexes formed with the E2 wild type and with the E2 peptide 3L, and 20 for the rest. Five extra simulations of 20.1 ns for the $E1^{WT}/E2^{WT}$ and three for $E1^{4L}/E2^{WT}$ were also performed to check the convergence of the trajectories. The results of the longer simulations were in perfect agreement with those of the 2.1 simulations (see Fig. 4). The final conformations were not constrained by the initial coordinates as was manifested by a wide range of very different conformations and two significant facts: 1), Most of the complexes found the correct left-handed packing even though the initial helix coordinates were chosen to be those of GpA, that is, initially the helices packed in a right-handed fashion. 2), Though all simulations were started with the parallel orientation existing in the virus (both C-terminals close to the viral nucleocapsid), some of them flipped over during the dynamics run resulting in antiparallel association, thus the helices were free to swivel up to 180°.

The simulations that exhibited antiparallel association were rejected for further analysis. In no case the $E1^{WT}/E2^{WT}$ complex and the $E1^{WT}/E2^{AN}$ complex formed those antiparallel arrangements, only the complexes $E1^{4L}/E2^{AN}$ and $E1^{11L}/E2^{AN}$ did so with six and four antiparallel conformations, respectively. Since the dipolar interaction between the helical dipoles is more favorable when the TM helices associate in an antiparallel fashion (Popot and Engelman, 1990), it is reasonable to think that the complexes that associated in that way were not provided with enough van der Waals interactions that could overcome the electrostatic antiparallel tendency. If the simulations were carried out with an explicit lipid environment, the lipid hydrocarbon chains would not allow such a flip and probably the helices had to stay apart from each other to get their equilibrium. Thus, even though some of these mutated

complexes can form virus particles in vivo, we can conclude that it is difficult for them to find the adequate contacts leading to the correct stable assembly, which is in agreement with the observed heterodimer and budding deficiencies of these mutations (Sjöberg and Garoff, 2003).

Analysis of the helical assembly using probability Cα-Cα contact matrices

In the wild-type and $E1^{WT}/E2^{AN}$ complexes simulations, most of the C_α-C_α contacts displayed in the C_α contact maps lay close to the diagonal, that is, the helices E1 and E2 packed their respective N- and C-terminals together, which is consistent with the length of the helices and the thickness of the lipid bilayer.

Probability contact matrices showing probabilities bigger than 10% were generated from the C_α contact maps (Fig. 3).

The probability contact matrices reflected significant differences between the different complexes, as well as an extreme similarity between the wild-type complex and the $E1^{WT}/E2^{AN}$ complex. In all complexes some residues of the E2 peptide did not make any contact with any of the residues of E1, horizontal empty lines appeared in the plots, consequently these residues faced the lipid, at least during the last 1.1 ns of simulation. These horizontal empty lines appeared in the plots evidenced a periodicity every three or four residues that clearly reflects the helical character of the E2 peptide. The E2 residues that faced the lipid were: 3, 7, 10, 11, 14, 15, 19, and 22. Surprisingly, all complexes displayed these residues at the lipid-helix interface. In contrast E1 did not show as clearly as E2 the residues facing the lipid. The wild-type as well as the $E1^{4L}/E2^{WT}$ complex and their respective analogs (the $E1^{4L}/E2^{AN}$ data is not shown) did not display almost any contact at residues Ile-10, Ile-13, and Val-17 of E1. The wild-type complex and its analog only differed in the third residue of the E2 peptide, a valine for the former and an alanine for the second (the same is applicable to $E1^{4L}/E2^{WT}$). Since the residue number 3 faces the lipid no differences are expected between the complexes involving whether $E2^{WT}$ or $E2^{AN}$. Their respective probability contact maps displayed the same contacts. A previous study of several GpA mutants proved that the mutations to alanine at sites facing the lipid do not affect significantly the free energy of helical association (Fleming and Engelman, 2001). Furthermore, they demonstrate that the dimer stability for different sequence variants in diverse hydrophobic environments conserves the

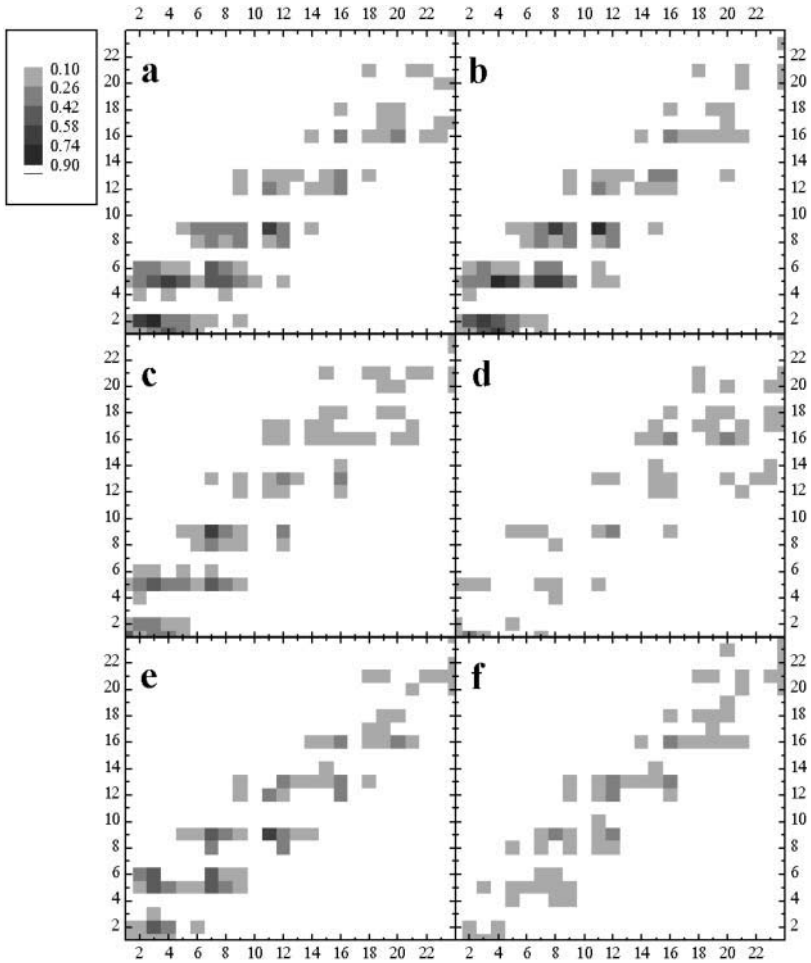

FIGURE 3 Probability contact matrices. Probability contact matrix between C_α atoms of E1 and E2 of the complexes: (*a*) $E1^{WT}/E2^{WT}$, (*b*) $E1^{WT}/E2^{AN}$, (*c*) $E1^{4L}/E2^{WT}$, (*d*) $E1^{11L}/E2^{WT}$, (*e*) $E1^{WT}/E2^{3L}$, and (*f*) $E1^{4L}/E2^{3L}$. The *x* axis accounts for the residue number of the helix E1 and the *y* axis for those of helix E2; the color reflects the normalized occurrence of each contact. The probabilities vary between 0.1 (*white*) and 0.9 (*black*). Contacts with probability lower than 0.1 were not plotted.

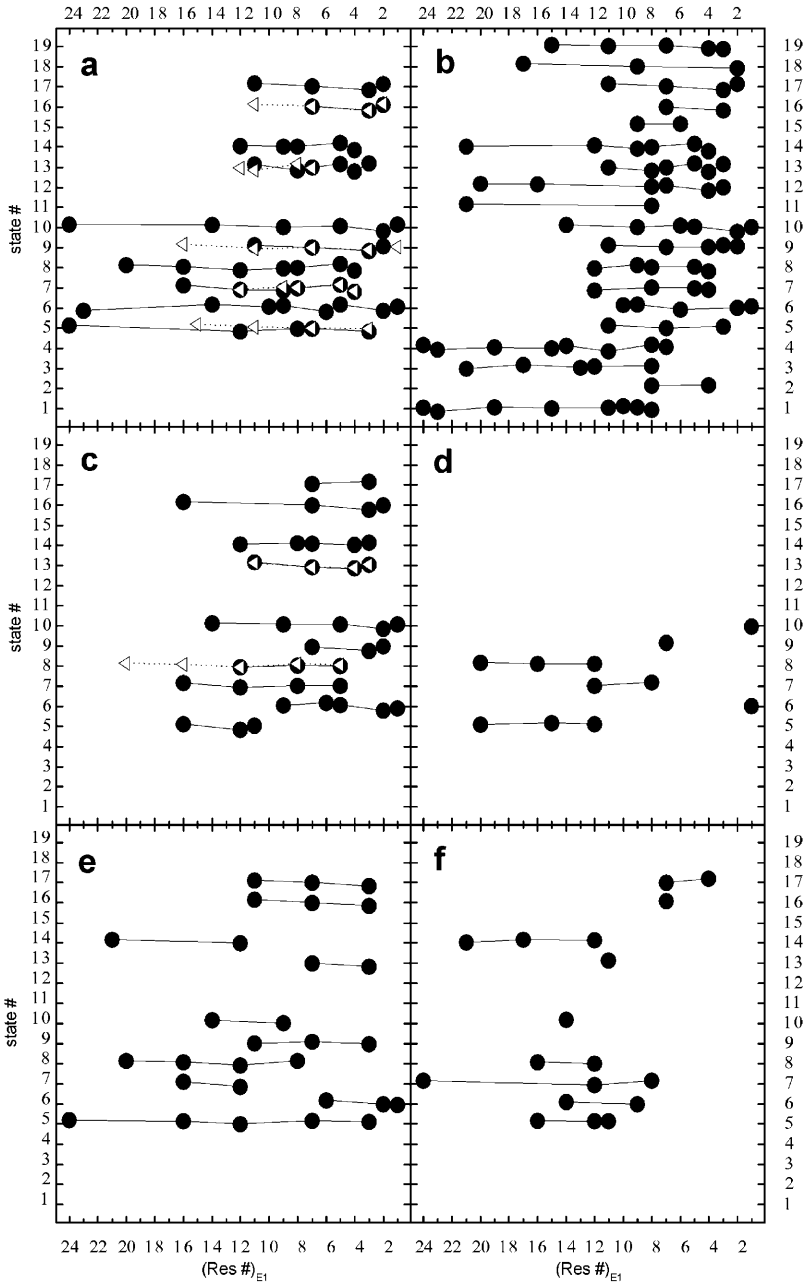

FIGURE 4 Helix packing curves. Helix packing curves of C_α(E1)-C_α(E2) distances of the E1/E2 complexes: (*a*) and (*a′*) E 1^{WT}/$E2^{WT}$, (*b*) and (*b′*) $E1^{WT}$/$E2^{AN}$, (*c*) and (*c′*) $E1^{4L}$/$E2^{WT}$, (*d*) and (*d′*) $E1^{11L}$/$E2^{WT}$, (*e*) and (*e′*) $E1^{WT}$/$E2^{3L}$, and (*f*) and (*f′*) $E1^{4L}$/$E2^{3L}$. The nonprimed cases account for the minimal distances from every E1 C_α atom to another E2 C_α atom and the primed case for the minimal distances from every E2 C_α atom to another E1 C_α atom. For each trajectory (*state #*) a symbol (*solid circle*, 2.1-ns simulations; *open triangle*, 20.1-ns simulations) is plotted only if the average C_α-C_α distance (averaged during the last 1.1 ns or 18 ns) was less than 6 Å. The vertical variation in the individual curves reflects the actual measured average distances, which were in the range 3.8–6 Å.

hierarchy of stability between the mutants. Since mutations at sites away from the helical interface do not influence the helix-helix packing, the complexes with the E2 wild type or its analog can be seen as equivalents. A corroboration of these ideas was obtained measuring the average total energy between all the simulations of the complexes $E1^{WT}$/$E2^{WT}$ and $E1^{WT}$/$E2^{AN}$, once the equilibrium was reached. In both cases the mean total energy was 1046 kcal/mol.

A cluster of contacts with high occurrence was concentrated in the N-terminal half of both helices in the wild-type and the $E1^{WT}$/$E2^{AN}$ complexes, whereas the regions closest to the C-terminal made scarcely any contacts, with typical distances larger than 12 Å. Poor contacts are shown in the complexes $E1^{11L}$/$E2^{WT}$ and $E1^{11L}$/$E2^{AN}$. In particular, the results revealed that these complexes lost the contacts at the 12 first residues from the N-terminal of both helices. In fact, the 12 leucines of the mutated complexes induced sterical clashes acting as a lever pushing apart the N-terminals of the helices. In some degree, the same occurred for the complexes containing the E1 peptide 4L. These peptides contained four leucines, residues 3–6, and one more leucine residue at position 11 where the E1 wild-type peptide had a glycine. However, those complexes were able to form contacts that were not as abundant as for the wild-type and the $E1^{WT}$/$E2^{AN}$ complexes but that showed a high level of occurrence. Such packing deficiencies correctly predict the defects of the

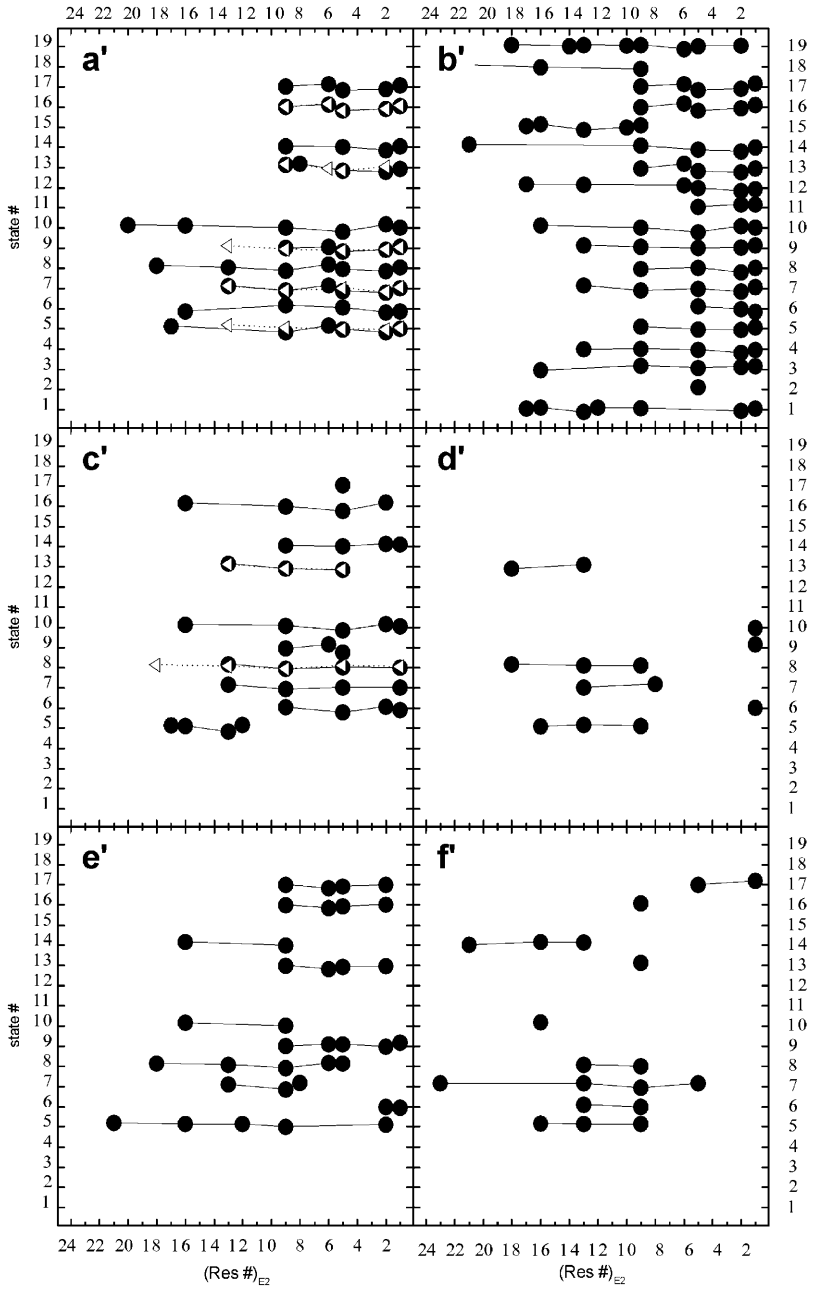

FIGURE 4 Continued.

mutated $E1^{4L}/E2^{WT}$ virus that form 40% less particles than the wild type. In summary, the E1 wild type offered a wider possible packing surface to E2 than the rest of complexes, that is, the E2 peptide had more possible candidate residues from E1 to get close to in the wild type than in the rest of complexes.

Analysis of the helical assembly using helix packing curves

The mean distances between C_α atoms of E1 and E2 of each one of the 10 states of the complexes $E1^{WT}/E2^{WT}$, $E1^{WT}/E2^{AN}$, $E1^{4L}/E2^{WT}$, $E1^{11L}/E2^{WT}$, $E1^{WT}/E2^{3L}$, and $E1^{4L}/E2^{3L}$ were calculated over the last 1.1 ns for the simulations of 2.1 ns and over the whole trajectory for the 20.1-ns simulations. Fig. 4 shows the residues for which the distance to some C_α atom of the opposite helix was less than 6 Å, together with their minimal distance. These curves inform clearly about the residues located at the interhelical interfaces, and one also sees that the same residues are located at the helical interfaces after 2.1 and 20.1 ns. Similar results were obtained for the 1.5-ns simulation of the $E1^{WT}/E2^{WT}$ complex with cutoff 30 Å (data not shown).

The helix packing curves of all simulations of the $E1^{WT}/E2^{WT}$ complex displayed the same periodic pattern: Ser-1, Ala-2, Gly-5, Met-6, and Leu-9 of the E2 peptide were in all cases in contact with some residue of peptide E1. Two other E2 residues, Ile-12 and Ser-13, made frequent contacts with

E1. Thus the part of the E2 peptide that was in contact with E1 exhibited the typical heptad motif, *abcdefg*, of the coiled-coils (Lupas, 1996) (see Table 3). In fact, this periodic contact motif, also known as leucine zipper, is just a consequence of the α-helical character of the peptides that form the coils when they pack together. Crick proposed in 1953 (Crick, 1953b) that α-helices tend to pack side by side in a knob-into-holes fashion if the helices coil around each other and are inclined at an appropriate angle (Crick, 1953a; Fraser and MacRae, 1973).

The contacts between helices were lost toward the C-terminals, in agreement with the results of the contact maps and with the larger fluctuations at the helix ends observed in other helical bilayer-spanning polypeptides simulations (Vogel et al., 1988). The minimal C_α-C_α distances were most of the times at E2 residues number 2, 5, and 9 with distances to some E1 C_α atom between 4 and 5 Å. In particular, the C_α atoms of the Ala-2 and Gly-5 residues were very often the closest to E1. Similarly the E1/E2 complexes $E1^{4L}/E2^{WT}$ and $E1^{4L}/E2^{AN}$ (data not shown) also exhibit the same E2 residues at the helical interface. In contrast the E2 peptide of the complex $E1^{11L}/E2^{WT}$ presented difficulties to make contacts to the E1 peptide: from the 10 simulations performed in this complex only four showed significant contacts but they did not preserve the same pattern as the $E1^{WT}/E2^{WT}$ complex. Moreover the contacts were achieved at the middle of the E2 helix and not close to the N-terminal.

The E1 helix contact pattern was not as well defined as for the E2 helix. Three different patterns were found for the $E1^{WT}/E2^{WT}$ and the $E1^{WT}/E2^{AN}$ complexes and are summarized in Table 3. An obvious four-periodicity was present in these patterns. Closer inspection of each individual case manifested that in fact the E1 helix packing patterns could also be explained by a knobs-into-holes heptad motif. The most common pattern presented residues Gly-3, Ala-7, and Gly-11 as contacting residues to E2 (case A of Table 3). For the second pattern, the residues at the E1 interhelical interface were Gly-4, Phe-8, Ala-9, and Ala-12, and in some cases also Leu-16 made close contact with E2 (case B of Table 3). The last and less common pattern had the C_α atoms of residues Ile-1, Ser-2, Leu-5, and Ala-9 as the closest atoms to E2 (case C of Table 3). In Fig. 5 is illustrated the first pattern (case A of Table 3). This figure shows how two ideal α-helices, E1 and E2, could achieve the knob-hole packing (*abcdefg* heptad motif) at the first turns of the helices if the two helices were mutually inclined in a right handed sense and slightly deformed to coil around each other forming a two-strand rope. In the figure, the E1 residues 3, 4, 7, 8, and 11 are at the interface. The second pattern (case B) can be reached from the first one (case A) by a rotation of $\sim 100°$, a rotation of $-100°$ around the E1 axis will put residues 4, 5, 8, 9, and 12 at the interhelical interface. Similarly, the third (case C) would be reached by a 150° rotation.

The two first packing patterns of the $E1^{WT}/E2^{WT}$ complex, which are the most favorable, are illustrated in Fig. 6. The helical wheels are shown with respect to the hyperhelical coil axis, assuming a perfect coiled-coil.

Also, Fig. 7 shows some snapshots of conformations that illustrate the three packing patterns. In the two first cases (Fig. 7, *a* and *b*) the helices packed in a left-handed two-stranded rope fashion with a interhelical dihedral angle of $-20°$. The inter N- and C-terminals distances were ~ 7 Å and 10 Å, respectively. This agreed with the larger distance between helices at the C-terminals than at the N-terminals observed by cryo-electron microscopy (Mancini et al., 2000). For the third pattern, illustrated in Fig. 7 *c*, the helices tended to pack with an interhelical axis angle close to 0°. It is clear that the helices got apart from each other as a result of the intrusion of the Leu-5 at the interhelical face. The distances between both N-terminals and both C-terminals in this case were 9 Å and 11 Å, respectively.

With respect to the $E1^{4L}/E2^{WT}$ complex, it displayed similar packing pattern as the $E1^{WT}/E2^{WT}$ complex, but in this case the leucine residues 4, 3, and 7 caused less close contacts to E2 than the wild type did and the interhelical axis angle was decreased (Fig. 7 *d*). However assuming that the complexes behave in a pure left-handed coiled-coil fashion, it can be seen that if the $E1^{4L}/E2^{WT}$ complex adopts the conformation shown in Fig. 6 *b*. For this complex only two mutated residues instead of three (Fig. 6 *a*) will be located at the interhelical interface, and they will be far away from the contact *a-d* dyad parallel to the hypercoil axis, affecting slightly the interfacial contacts. The heterodimer stability of the virus will be affected since the contacts between helices are weaker even though it presents approximately the same interhelical crossing area.

The E1 helix of the $E1^{11L}/E2^{WT}$ complex occasionally made contacts with the E2 peptide. When contacts occurred

TABLE 3 Characteristic heptad motif $(abcdefg)_n$ of the knobs-into-holes packing

		g	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*a*	*b*	*c*	*d*	*e*
E2		1	2	3	4	5	6	7	8	9	10	11	12	13
		Ser	Ala			Gly	Met			Leu			Ile	Ser
E1	A	3	4	5	6	7	8	9	10	11	12	13	14	15
		Gly	Gly			Ala	Phe			Gly				
	B	4	5	6	7	8	9	10	11	12	13	14	15	16
		Gly	Leu			Phe	Ala			Ala				Leu
	C	1	2	3	4	5	6	7	8	9	10	11	12	13
		Ile	Ser			Leu	Gly			Ala				

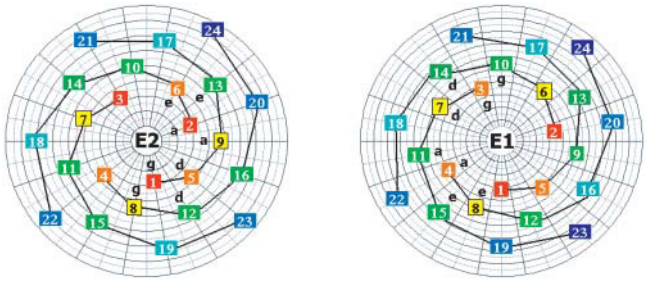

FIGURE 5 α-Helical spiral wheels. Helical spiral representation of the ideal α-helices of the E1/E2 wild-type complex respect to their own helical axis. The Cα position of each residue is projected along the helical axis. The view is from the C-terminal toward the N-terminal. Red corresponds to the closest turn to the N-terminal and the dark blue to the last turn at the C-terminal. Every Cα^{i} atom is bound to the closest Cα atom, that is to the Cα^{i+4}. The positions *a*, *d*, *e*, and *g* of the typical coiled-coil heptad motif are also marked.

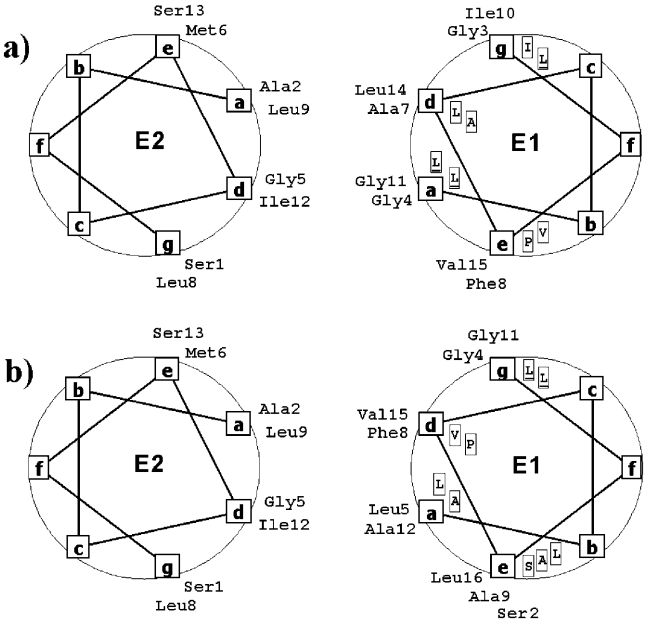

FIGURE 6 Two-strand rope helical wheels. Helical wheel with respect to the superhelical coil axis: The E1 and E2 wild type are shown as perfect coiled-coils with periodicity 3.5 with respect to the supercoil axis. The first figure (*a*) represents case A of Table 3, whereas the second (*b*) accounts for case B. The residues on the squares inside the circle of the E1 helical wheel are the corresponding residues of the 4L peptide in the case that $E1^{4L}/E2^{WT}$ complex will adopt the same disposition as the wild-type complex. The underlined residues are those that are mutated in 4L respect to the E1 wild type. In *a*, the mutated leucines at position 4 and 11 that are glycine in the wild-type peptide will be at the helical interface at the *a* position of the *abcdefg* heptad, loosing the close packing that the glycines provided. In *b*, the mutations occupy *g* positions and the disruption of the interhelical packing will be smaller than in *a*.

they appeared at the end of both helices with larger C_α-C_α distances than those of the wild-type complex. Fig. 7 *e* represents a snapshot of one of the $E1^{11L}/E2^{WT}$ conformations that possessed the largest contact region among the 10 simulations of this complex. The contacts were completely lost toward the N-terminals and the 12 leucines of E1 tended to bind to the leucines of E2 disrupting the contact motif present in the wild-type complex (the residues Ser-1, Ala-2, Gly-5, Leu-9, Ile-12, and Ser-13 were not longer at the interhelical interface of E2). In this case the bulky leucine residues of E1 did not find in E2 a suitable complementary surface to pack against to, in contrast with the soluble leucine zippers where the hydrophobic effect drives the close packing. Empty spaces were created between both helices. In addition, the interhelical axis angle was also disrupted with a clear tendency to a parallel orientation of the helices.

Hydrogen bonds

A small number of interhelical hydrogen bonds were formed. These were very often found between the first residues close to the N-terminals of the helices of the wild-type complex and less frequently in the $E1^{4L}/E2^{WT}$ complex, whereas interhelical hydrogen bonds of the $E1^{11L}/E2^{WT}$ complex occurred mainly close to the C-terminals.

Analysis of the intrahelical hydrogen bonds manifested that as in the single helix simulations the helices E1 and E2 in complex had also a π-helical tendency. Segments of the helices became π-helical when the contacts with the other helix were lost. The wild-type complex, its analog, and the $E1^{4L}/E2^{WT}$ complex presented π-helical conformation at the last third of both helices toward the C-terminals. The contacts in the $E1^{11L}/E2^{WT}$ complex were mostly localized at the middle toward the end of the helices and consequently the π-helical regions were shorter than in the previous cases.

A special case was the E1 residue Gly-3 of the wild-type complex; this amino acid did not form intrahelical hydrogen bonds. The effect of this can be the formation of a kink in the E1 helix at position Ala-7, in concordance with the results of the single helix simulations that showed how the E1 wild type was much more flexible than the rest of E1 peptides. Furthermore glycines can adopt unusual dihedral angles perturbing the helical backbone, as is clearly exhibited in Fig. 7 *b*. As mentioned before, inside the lipid membrane the C_α-H$\cdots$O hydrogen bonds can be important. When the C_α-H$\cdots$O bonds were studied we found that the Gly-3 carbonyl oxygen bonded to the α-hydrogen of Ala-2 up to 25% of the time in some simulations and also to the α-hydrogen of Ser-1, preventing α-helical hydrogen bond formation. Furthermore, the Gly-3 α-hydrogen bonded to the Ala-2 carbonyl oxygen up to 40% of the time in some simulations. Similarly the Gly-4 carbonyl oxygen was bonded to the α-hydrogen of Gly-5 and/or Ala-2. A statistical analysis of the C_α-H$\cdots$O hydrogen bonds over all the simulations of every complex revealed that in the wild-type complex these bonds formed very often close to the N-terminal; the C_α-H$\cdots$O hydrogen bonds that formed most often in E1-E2 were: $(Ser\text{-}2)^{E1}$–$(Ala\text{-}2)^{E2}$, 8.2%; $(Gly\text{-}3)^{E1}$–$(Ser\text{-}1)^{E2}$, 11.7%; $(Gly\text{-}3)^{E1}$–$(Ala\text{-}2)^{E2}$, 7.7%; and $(Gly\text{-}4)^{E1}$–$(Gly\text{-}5)^{E2}$, 10.5%. The $E1^{4L}/E2^{WT}$ complex presented analogous bonds, in particular the

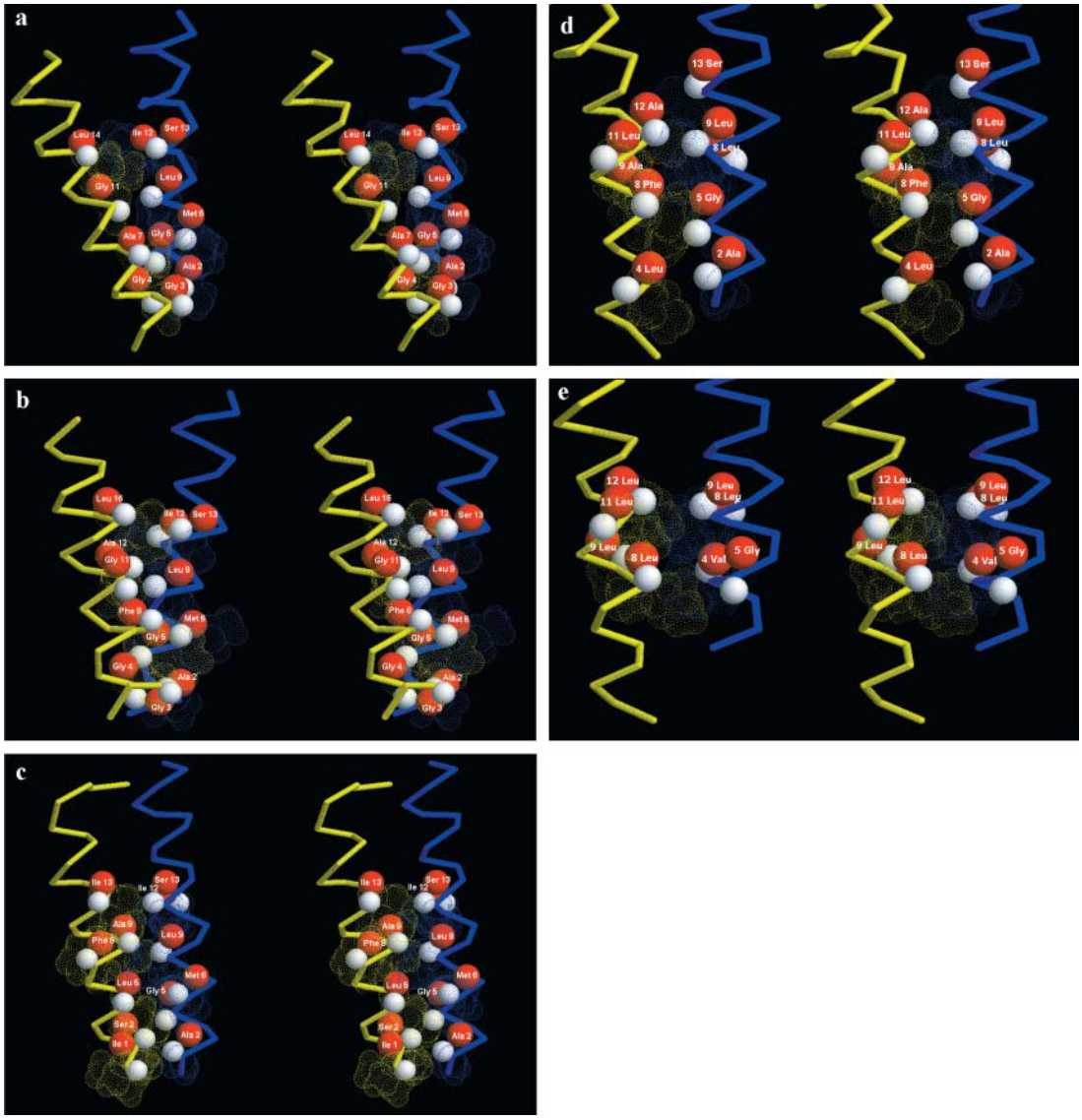

FIGURE 7 Snapshots. Stereo view of some snapshots of complexes E1/E2 at 2.1 ns. The three first figures correspond to different simulations of the $E1^{WT}$/$E2^{WT}$ complex. The E1 residues at the interhelical interface correspond to (*a*) a typical A case of Table 3, (*b*) a B case, and (*c*) a C case. In *a* and *b*, the helices formed a two-stranded rope, whereas in *c* the helices loosed practically the left-handed fashion of the rope. Snapshots of $E1^{4L}$/$E2^{WT}$ and $E1^{11L}$/$E2^{WT}$ are shown in *d* and *e*, respectively. The substitutions to leucines hindered the close contacts at the N-terminals; the effect is stronger for the $E1^{11L}$ peptide. The helices got more parallel to each other.

E1 Ser-2 and Leu-4 residues bound with some frequency to the E2 residues Ala-2 and Gly-5. In summary, both complexes form a small network of C_α-H$\cdots$O hydrogen bonds between the two helices close to their N-terminal keeping them intimately packed. In contrast to this, the $E1^{11L}$/$E2^{WT}$ complex more rarely formed this kind of hydrogen bonds and when they occurred they appeared close to the C-terminal.

Complexes with the $E2^{3L}$ peptide

In the virus the $E1^{4L}$/$E2^{WT}$ complex forms a more unstable E1/E2 heterodimer and reduces particle formation to ~60% of that of the wild type (Sjöberg and Garoff, 2003). From our simulations an explanation for this behavior can be given: the $E1^{4L}$/$E2^{WT}$ complex can conserve roughly the wild-type complex helical packing. If the E1 and E2 helices of the mutant were able to find the right orientation to pack according to the models proposed in Fig. 6, only two or three leucines of the mutated residues would be at the interhelical interface. In particular, if the conformation shown in Fig. 6 *b* was reached, two of the point mutations will happen at position *g* at the lateral side of the contact *a*-*d* dyad, keeping then the same crossing area structure as the wild type, whereas the other two mutated residues will be located at the helical-lipid interface having a small effect on the helix-helix packing as explained previously.

The E2 peptide of this complex aimed the same residues at the interhelical interface as the wild-type complex. That is, the interhelical surface of the wild-type E2 peptide was complementary to the interhelical surface of the 4L peptide, even though not so good contacts were accomplished as in the wild-type complex case. In conclusion, the success of the $E1^{4L}/E2^{WT}$ packing is a compromise between both E1 and E2 peptides.

To test this theory we constructed a new E2 peptide called 3L (see Tables 1 and 2). In this case the residues 2 and 5, located at the E2 surface of the wild-type complex, were mutated to leucines. The peptide carried two more mutations, Val-3/Ala and Ala-10/Leu, which a priori should face the lipid face. We aspired to generate a complex with the 4L peptide that will not preserve the packing between helices of the $E1^{4L}/E2^{WT}$ complex, although in complex with the E1 wild-type peptide should nearly preserve the wild-type complex packing at least as the $E1^{4L}/E2^{WT}$ complex did.

We performed the same analysis procedures as before with the two new complexes, $E1^{WT}/E2^{3L}$ and $E1^{4L}/E2^{3L}$. The probability contact map and the helix packing curves (Figs. 3, *e* and *f*, and 4, *e*, *e′*, *f*, and *f′*) are consistent with our hypothesis. The $E1^{WT}/E2^{3L}$ complex made contacts close to the N-terminal of both helices although the contacts are lost at the very beginning. The E2 contact surface was made up of residues Leu-2, Leu-5, and Leu-9, but Ser-1 lost its close packing, whereas the E1 peptide preserved the Gly-3, Ala-7, and Gly-11 packing pattern most of the time. In the $E1^{4L}/E2^{3L}$ complex case the contacts were shifted toward the middle of the helices, and the E2 peptide lost again as for the $E1^{11L}/E2^{WT}$ complex its typical contact residues at the interhelical surface. The analysis of the occurrence C_α-H···O hydrogen bonds evidenced that for the former case the E1 wild-type Gly-3 was able to form C_α-H···O hydrogen bonds frequently to Leu-2 of the 3L peptide. The $E1^{4L}/E2^{3L}$ complex rarely formed C_α-H···O hydrogen bonds, and when it happened it was close to the C-terminals of both helices.

Implications for other related viruses

The sequences of the TM domains within a virus family are largely conserved and consequently their structures are similar. The alphaviruses present a high sequence similarity for the E1 peptide. The E2 peptide on the contrary does not reveal such a high similarity. Table 4 shows the sequences of the E1 and E2 putative TM spanning helices of the Sindbis virus, the Ross River virus, and the Western and Venezuelan Equine Encephalitis virus (WEE (Hahn et al., 1988) and VEE, respectively) that belong to the alphavirus genus of the *Togaviridae* family. Two glycine residues of the TM spanning domain of E1 are clearly conserved among these viruses, whereas no obvious conservation occurred for E2.

Assuming that these viruses adopted the left-handed coiled-coil structure of the SFV, the residues that all of them will present at the putative interhelical positions 3, 4, 7, 8, 9, 11, and 12 for E1 and 1, 2, 5, 6, 9, 12, and 13 for E2 (positions *g*, *a*, *d*, *e*, *g*, and *a* of the heptad motif) were of the same kind. These residues were small amino acids as glycines, alanines, and serines and bulkier hydrophobic amino acids as leucines, isoleucines, and valines. Thus, even though the putative residues at the interhelical interfaces are not exactly the same, the kind of contacts that E1 and E2 will establish is the same. This agrees with the idea of the E1-E2 compromise for correct packing. We could observe that when bulky amino acids that could hinder close packing were present in E1, E2 compensates this effect by introducing smaller residues in the interhelical face to which the former could pack against. For instance, the E1 peptide of the Sindbis virus has a serine and two leucines at positions 7, 8, and 9 that hypothetically should pack between residues 5, 6, and 9 of the E2 peptide. These three positions are in this case occupied by a serine and alanines, whereas in the E2 SFV small residues were not so necessary at these positions because the interhelical residues in E1 were small.

The glycine, alanine, and serine amino acids that occupy the putative positions of the interhelical interface not only favor the close packing but also provide a suitable scaffold for the formation of C_α-H···O hydrogen bond networks.

TABLE 4 Sequences of related alphaviruses

E1WT	SFV4	ISGGLGAFAIGAILVLVVVTCIGL
	RossR	MASGLGGLALIAVVVLVLVTCITM
	Sin	LFGGASSLLIIGLMIFACSMMLTS
	WEE	LFGGASSLIVVGLIVLVCSSMLIN
	VEE	LLGGSAVIIIIGLVLATIVAMYVL
E2WT	SFV	SAVVGMSLLALISIFASCYMLVAA
	RossR	AAVSGASLMALLTLAATCCMLATA
	Sin	LAVASATVAMMIGVTVAVLCACKA
	WEE	IVLCGVALAILVGTASSAACIAKA
	VEE	LGLSICAAIATVSVAASTWLFCRS

CONCLUSION

The results from the simulations suggest that the E1/E2 complex of the SFV TM spanning domain is a complex of two helices that intimately pack against each other close to the N-terminals in a parallel left-handed two-stranded rope fashion. The success of this packing is a compromise of small and medium complementary residues at both interhelical interfaces (Gly, Ala, Ser, and Leu) that facilitate the formation of a network of C_α-H···O hydrogen bonds between the E1 and E2 peptides.

Since high sequence identity is in most cases equivalent to similar structure, the SFV related viruses may possess the heptad motif characteristic of the coiled-coils hidden inside their TM spanning domain. In Table 4 are listed the possible residues at the interhelical interfaces of some these. The apparent nonconservation of several residues can be understood in terms of the previously referred E1-E2

compromise: a substitution at the E1 interhelical interface implies a change in the E2 peptide that compensates the effect induced by E1.

With respect to the simulations per se, it might be stated that 2.1 ns is a very short time to insure convergence. However, as argued above, the results from longer simulations illustrated in Fig. 4 demonstrated that 2.1 ns of simulation is enough for our purposes. Given that the lipid bilayer is not a uniform medium (Woolf, 1998) and the interactions lipid/protein play an important role in the structure, stability, and dynamics of TM α-helices (Petrache et al., 2000), the use of explicit lipids instead of a continuum dielectric medium and longer simulations would give a more accurate description of the helix-helix interaction. In this work we have not intended to give a detailed atomic description of the helical packing but to localize the residues confined to the helical interfaces that can explain the E1/E2 heterodimer stability of the SFV alphavirus in the membrane-spanning domain. The strategy that we have developed can be applied to other alphaviruses as an initial approach to characterize their TM structure and as a tool for the design of new mutants.

We thank Mathilda Sjöberg and Jan Norberg for stimulating discussions.

We also thank the Swedish Research Council and the National Graduate School for Scientific Computing for financial support.

REFERENCES

Adams, P. D., D. M. Engelman, and A. T. Brunger. 1996. Improved prediction for the structure of the dimeric transmembrane of glycophorin A obtained through global searching. *Proteins.* 26:257–261.

Bright, J. N., and M. S. P. Sansom. 2003. The flexing/twirling helix: exploring the flexibility about molecular hinges formed by proline and glycine motifs in transmembrane helices. *J. Phys. Chem. B.* 107: 627–636.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Crick, F. H. 1953a. The Fourier transform of a coiled-coil. *Acta Crystallogr.* 6:685–689.

Crick, F. H. 1953b. The packing of alpha-helices: simple coiled-coils. *Acta Crystallogr.* 6:689–697.

Duneau, J. P., S. Crouzy, N. Garnier, Y. Chapron, and M. Genest. 1999. Molecular dynamics simulations of the ErbB-2 transmembrane domain within an explicit membrane environment: comparison with vacuum simulations. *Biophys. Chem.* 76:35–53.

Duneau, J. P., D. Genest, and M. Genest. 1996. Detailed description of an alpha helix→pi bulge transition detected by molecular dynamics simulations of the p185c-erbB2 V659G transmembrane domain. *J. Biomol. Struct. Dyn.* 13:753–769.

Eilers, M., S. C. Shekar, T. Shieh, S. O. Smith, and P. J. Fleming. 2000. Internal packing of helical membrane proteins. *Proc. Natl. Acad. Sci. USA.* 97:5796–5801.

Feig, M., A. D. MacKerell, and C. L. Brooks. 2003. Force field influence on the observation of pi-helical protein structures in molecular dynamics simulations. *J. Phys. Chem. B.* 107:2831–2836.

Fleming, K. G., and D. M. Engelman. 2001. Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc. Natl. Acad. Sci. USA.* 98:14340–14344.

Forrest, L. R., A. Kukol, I. T. Arkin, D. P. Tieleman, and M. S. Sansom. 2000. Exploring models of the influenza A M2 channel: MD simulations in a phospholipid bilayer. *Biophys. J.* 78:55–69.

Fraser, R. D. B., and T. P. MacRae. 1973. Conformation in Fibrous Proteins and Related Synthetic Polypeptides. Academic Press: New York and London.

Garoff, H., and R. H. Cheng. 2001. The missing link between envelope formation and fusion in alphaviruses. *Trends Microbiol.* 9:408–410.

Haag, L., H. Garoff, L. Xing, L. Hammar, S. T. Kan, and R. H. Cheng. 2002. Acid-induced movements in the glycoprotein shell of an alphavirus turn the spikes into membrane fusion mode. *EMBO J.* 21:4402–4410.

Hahn, C. S., S. Lustig, E. G. Strauss, and J. H. Strauss. 1988. Western equine encephalitis virus is a recombinant virus. *Proc. Natl. Acad. Sci. USA.* 85:5997–6001.

Henderson, R. 1975. The structure of the purple membrane from Halobacterium hallobium: analysis of the X-ray diffraction pattern. *J. Mol. Biol.* 93:123–138.

Henderson, R. 1977. The purple membrane from Halobacterium halobium. *Annu. Rev. Biophys. Bioeng.* 6:87–109.

Javadpour, M. M., M. Eilers, M. Groesbeek, and S. O. Smith. 1999. Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J.* 77:1609–1618.

Keefe, L. J., J. Sondek, D. Shortle, and E. E. Lattman. 1993. The alpha aneurism: a structural motif revealed in an insertion mutant of staphylococcal nuclease. *Proc. Natl. Acad. Sci. USA.* 90:3275–3279.

Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567–580.

Laage, R., J. Rohde, B. Brosig, and D. Langosch. 2000. A conserved membrane-spanning amino acid motif drives homomeric and supports heteromeric assembly of presynaptic SNARE proteins. *J. Biol. Chem.* 275:17481–17487.

Law, R. J., L. R. Forrest, K. M. Ranatunga, P. La Rocca, D. P. Tieleman, and M. S. Sansom. 2000. Structure and dynamics of the pore-lining helix of the nicotinic receptor: MD simulations in water, lipid bilayers, and transbilayer bundles. *Proteins.* 39:47–55.

Lee, K. H., D. R. Benson, and K. Kuczera. 2000. Transitions from alpha to pi helix observed in molecular dynamics simulations of synthetic peptides. *Biochemistry.* 39:13737–13747.

Lemmon, M. A., H. R. Treutlein, P. D. Adams, A. T. Brunger, and D. M. Engelman. 1994. A dimerization motif for transmembrane alpha-helices. *Nat. Struct. Biol.* 1:157–163.

Liu, L. P., and C. M. Deber. 1998. Uncoupling hydrophobicity and helicity in transmembrane segments. Alpha-helical propensities of the amino acids in non-polar environments. *J. Biol. Chem.* 273:23645–23648.

Lupas, A. 1996. Coiled coils: new structures and new functions. *Trends Biochem. Sci.* 21:375–382.

MacKenzie, K. R., J. H. Prestegard, and D. M. Engelman. 1997. A transmembrane helix dimer: structure and implications. *Science.* 276:131–133.

MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102: 3586–3616.

Mancini, E. J., M. Clarke, B. E. Gowen, T. Rutten, and S. D. Fuller. 2000. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol. Cell.* 5:255–266.

Morgan, D. M., D. G. Lynn, H. Miller-Auer, and S. C. Meredith. 2001. A designed Zn2+-binding amphiphilic polypeptide: energetic consequences of pi-helicity. *Biochemistry*. 40:14020–14029.

Norberg, J., and L. Nilsson. 2000. On the truncation of long-range electrostatic interactions in DNA. *Biophys. J*. 79:1537–1553.

Pappu, R. V., G. R. Marshall, and J. W. Ponder. 1999. A potential smoothing algorithm accurately predicts transmembrane helix packing. *Nat. Struct. Biol*. 6:50–55.

Petrache, H. I., A. Grossfield, K. R. MacKenzie, D. M. Engelman, and T. B. Woolf. 2000. Modulation of glycophorin A transmembrane helix interactions by lipid bilayers: molecular dynamics calculations. *J. Mol. Biol*. 302:727–746.

Popot, J. L., and D. M. Engelman. 1990. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry*. 29:4031–4037.

Popot, J. L., and D. M. Engelman. 2000. Helical membrane protein folding, stability, and evolution. *Annu. Rev. Biochem*. 69:881–922.

Rajashankar, K. R., and S. Ramakumar. 1996. Pi-turns in proteins and peptides: classification, conformation, occurrence, hydration and sequence. *Protein Sci*. 5:932–946.

Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys*. 23: 327–341.

Scheiner, S., T. Kar, and Y. Gu. 2001. Strength of the Calpha H.O hydrogen bond of amino acid residues. *J. Biol. Chem*. 276:9832–9837.

Senes, A., I. Ubarretxena-Belandia, and D. M. Engelman. 2001. The Calpha–H.O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA*. 98: 9056–9061.

Singer, S. J. 1990. The structure and insertion of integral proteins in membranes. *Annu. Rev. Cell Biol*. 6:247–296.

Singer, S. J., and M. P. Yaffe. 1990. Embedded or not? Hydrophobic sequences and membranes. *Trends Biochem. Sci*. 15:369–373.

Sjöberg, M., and H. Garoff. 2003. Interactions between the transmembrane segments of the alphavirus E1 and E2 proteins play a role in virus budding and fusion. *J. Virol*. 77:3441–3450.

Sternberg, M. J., and W. J. Gullick. 1989. Neu receptor dimerization. *Nature*. 339:587.

Sternberg, M. J., and W. J. Gullick. 1990. A sequence motif in the transmembrane region of growth factor receptors with tyrosine kinase activity mediates dimerization. *Protein Eng*. 3:245–248.

Treutlein, H. R., M. A. Lemmon, D. M. Engelman, and A. T. Brunger. 1992. The glycophorin A transmembrane domain dimer: sequence-specific for a right-handed supercoil of helices. *Biochemistry*. 31: 12726–12732.

Vogel, H., L. Nilsson, R. Rigler, K. P. Voges, and G. Jung. 1988. Structural fluctuations of a helical polypeptide traversing a lipid bilayer. *Proc. Natl. Acad. Sci. USA*. 85:5067–5071.

Walther, D., F. Eisenhaber, and P. Argos. 1996. Principles of helix-helix packing in proteins: the helical lattice superposition model. *J. Mol. Biol*. 255:536–553.

Weaver, T. M. 2000. The pi-helix translates structure into function. *Protein Sci*. 9:201–206.

Whitley, P., L. Nilsson, and G. von Heijne. 1993. Three-dimensional model for the membrane domain of Escherichia coli leader peptidase based on disulfide mapping. *Biochemistry*. 32:8534–8539.

Woolf, T. B. 1997. Molecular dynamics of individual alpha-helices of bacteriorhodopsin in dimyristol phosphatidylocholine. I. Structure and dynamics. *Biophys. J*. 73:2376–2392.

Woolf, T. B. 1998. Molecular dynamics simulations of individual alpha-helices of bacteriorhodopsin in dimyristoylphosphatidylcholine. II. Interaction energy analysis. *Biophys. J*. 74:115–131.