Wolfgang B. Fischer
Lucy R. Forrest
Graham R. Smith
Mark S. P. Sansom
Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, Rex Richards Building, South Parks Road, Oxford, OX1 3QU, UK

Received 1 August 1999; accepted 17 November 1999

Transmembrane Domains of Viral Ion Channel Proteins: A Molecular Dynamics Simulation Study

Abstract: *Nanosecond molecular dynamics simulations in a fully solvated phospholipid bilayer have been performed on single transmembrane α-helices from three putative ion channel proteins encoded by viruses: NB (from influenza B), CM2 (from influenza C), and Vpu (from HIV-1). α-Helix stability is maintained within a core region of ca. 28 residues for each protein. Helix perturbations are due either to unfavorable interactions of hydrophobic residues with the lipid headgroups or to the need of the termini of short helices to extend into the surrounding interfacial environment in order to form H-bonds. The requirement of both ends of a helix to form favorable interactions with lipid headgroups and/or water may also lead to tilting and/or kinking of a transmembrane α-helix. Residues that are generally viewed as poor helix formers in aqueous solution (e.g., Gly, Ile, Val) do not destabilize helices, if located within a helix that spans a lipid bilayer. However, helix/bilayer mismatch such that a helix ends abruptly within the bilayer core destabilizes the end of the helix, especially in the presence of Gly and Ala residues. Hydrogen bonding of polar side-chains with the peptide backbone and with one another occurs when such residues are present within the bilayer core, thus minimizing the energetic cost of burying such side-chains.* © 2000 John Wiley & Sons, Inc. Biopoly 53: 529–538, 2000

Keywords: *molecular dynamics simulation; viral ion channel; transmembrane alpha-helix; alpha-helix propensity; lipid bilayer; hydrophobic mismatch; hydrogen bonding*

INTRODUCTION

The majority of integral membrane proteins contain transmembrane (TM) domains formed by bundles of α-helices. In particular, a bundle of α-helices packed together in a rotationally symmetrical fashion about a central axis coincident with a transbilayer pore is a structural motif common to a number of ion channels. For example, amphipathic α-helical peptides such as alamethicin[1,2] may generate channels by self-assem-

Correspondence to: Wolfgang B. Fischer; e-mail: wolfgang@bioch.ox.ac.uk
Contract grant sponsor: European Union (TMR)
Contract grant sponsor: Wellcome Trust
Biopolymers, Vol. 53, 529–538 (2000)

bly to form a simple α-helix bundle. More complex channel proteins with a pore domain based upon a TM helix bundle are known, such as the nicotinic acetylcholine receptor,[3] the bacterial K^+-channel KcsA[4] and the bacterial mechanosensitive channel MscL.[5] Thus, it is of some interest to understand the principles underlying the structure and stability of such assemblies of TM helices. A first step toward this requires an understanding of the factors stabilizing formation of isolated TM helices. Conceptually, such an approach follows from the 2-stage model for membrane protein folding first proposed by Popot and Engleman.[6] In this model, helices are inserted into a bilayer (stage 1) and form stable TM domains prior to subsequent self-assembly (stage 2) to form helix bundles that correspond to a folded membrane protein.

A relatively simple class of ion channels is provided by a number of virally encoded membrane proteins.[7,8] These may provide a paradigm for ion channel folding and assembly. The channels are formed by homo-oligomeric helix bundles, each subunit of the assembly providing a single TM helix. Thus, these proteins are sufficiently simple to be amenable to a range of biophysical techniques,[9–14] while providing models of arguably greater biomedically relevance than, e.g., membrane active peptides.[15]

The most intensively studied viral ion channel is M2 from influenza A, which forms proton selective pores activated by low pH.[16–18] M2 channels aid viral uncoating by facilitating acidification of the interior of the virion once within the endosome of an infected cell. M2 channels are blocked (i.e., inhibited) by the anti-influenza drug amantadine.[19,20] More recently, it has been shown that the viral genomes of influenza B and C also encode simple integral membrane proteins, NB (100 amino acids) and CM2 (115 amino acids), respectively. These proteins resemble M2 in that: (i) each contains a single TM helix close to its N terminus[21–23]; (ii) small numbers of copies of the proteins are present in the membrane of each parent virion; and (iii) the peptide chains homo-oligomerize to form the intact protein.[21,22,24,25] There are reasonable data to support the contention that NB and CM2 may form ion channels analogous to M2, although the roles of these putative channels within the life cycles of their respective viruses are incompletely understood.[21,24] NB exhibits complex conductance properties, i.e., it induces both an Na^+ conductance at neutral pH and a Cl^- conductance at low pH[26] in addition to an Na^+-dependent H^+ conductance.[27] CM2 resembles M2 in that it forms disulphide-linked dimers and tetramers on SDS-gels under nonreducing conditions.[25] Given these similarities, it is possible that CM2 may form ion channels.[22]

In addition to ion channels encoded by influenza viruses, a possible ion channel is found in HIV-1. The HIV-1 genome encodes Vpu, a 16 kDa phospho-protein of ca. 81 residues, which also contains a single TM helix near its N-terminus.[28] Vpu is expressed by HIV-1 infected cells, forming homo-oligomers in the cell membranes. Vpu appears to be a bifunctional protein.[29] Its TM domain plays a role in virus release/secretion,[30] whereas the extramembraneous C-terminal domain regulates CD4 degradation.[31,32] Electrophysiological experiments with Vpu expressed in E. coli and reconstituted into lipid bilayers[33] or Vpu exogenously expressed in Xenopus oocytes[34] suggest that a monovalent cation conductance may be induced. This is supported by planar bilayer studies of a synthetic peptide corresponding to the Vpu TM domain.[34] However, the exact nature and role of Vpu-induced channels remains to be fully demonstrated.[7,35]

Despite the topological and functional similarities of M2, NB, CM2, and Vpu, their sequences do not exhibit any detectable homologies with one another. Although such lack of homology does not preclude a similar TM structure, it compounds one of the problems of modeling and simulation studies of TM helices of these viral ion channels in the absence of definitive crystallographic or NMR structures. Although methods for prediction of the location of TM helices within the sequences of membrane proteins are well established, such methods suffer from a degree of inaccuracy in prediction of the exact start and end residues. As shown for, e.g., M2, most programs agree in predicting a core hydrophobic region.[36] However, for M2 this was, at 18 residues, too short to form a stable α-helical TM segment. MD simulations of M2 α-helical models of different lengths embedded in fully hydrated phospholipid bilayer aided in determination of the optimal length of a stable TM helix in its "native" environment. Thus, MD simulations may be used to refine sequence-based predictions of TM helices. In some ways, this parallels the use of extended MD simulations to study the folding behavior, in isotropic solvents, of peptides[37] and small proteins.[39]

In this article, we describe comparative MD simulations of the TM segments of NB, CM2, and Vpu. TM helix models of varying lengths are generated on the basis of several sequence-based TM prediction algorithms, and are used as starting structures for nanosecond duration MD simulations of the helices embedded in a phospholipid bilayer. The results of these simulations are used to assess the most stable length of each TM helix when in a phospholipid bilayer environment. Analysis of hydrogen bonding

patterns sheds light on protein/lipid/water interactions and their relationship to helix stability.

MATERIALS AND METHODS

Sequences of the three putative viral ion channels (NB, Lee; CM2, Q67389; and Vpu, HV1H2) were obtained from the SwissProt database (http://www.nick.med.usf.edu/GCGdoc/Data_Files/swissprot.html). The TM prediction programs used were MEMSAT v1.7 (http://globin.bio.warwick.ac.uk/~jones/memsat.html)[39]; DAS v2 (http://www.biokemi.su.se/~server/DAS)[40]; TMAP v1.4 (http://www.embl-heidelberg.de/tmap/tmap-info.html)[41]; TopPred2 (http://www.biokemi.su.se/~server/TopPred2)[42]; PHDhtm (http://www.embl-heidelberg.de/predictprotein)[43]; and TMHMM (http://www.cbs.dtu.dk/services/TMHMM-1.0/).

Initial models of TM α-helices were generated using Xplor[44] by restrained in vacuo MD simulations.[45] All side-chains were modeled as being in their default ionization state. Each helix thus generated was docked in a hole of radius 0.7 nm in a pre-equilibrated lipid bilayer (of 1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphatidylcholine (POPC)) of area ca. 6×6 nm,^{2} as described in detail elsewhere.[36,46] Each lipid/protein system was then solvated with more than 30 SPC water molecules per lipid. MD simulations were performed using GROMACS (http://rugmd0.chem.rug.nl/~gmx/gmx.html). All simulations were of duration 1 ns. Simulations were run on either a 10 or an 80 processor SGI Origin 2000. For a system of ca. 20,000 atoms, the cpu time is ca. 8 days per ns per processor. Subsequent analysis of the simulations used Gromacs and DSSP.[47] Structures were visualized using Quanta (Biosym/MSI) and Rasmol.

RESULTS AND DISCUSSION

TM-Predictions

All six sequence-based methods overlap in the prediction of a helical TM region extending from ca. residue 18 to residue 42 for NB (RGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KI, see also Fig. 1), which can be identified as a core region. The agreement among different methods for CM2 is poorer, but the most probable TM region lies between residues 57–80 (LASL^{60}GLGIITMLYL^{70}LVKIIIELVN^{80}). For Vpu, the core TM region is predicted to be between residues ca. 5–28 (IVAIVA^{10}LVVAIIIAIV^{20}VWSIVIIE). However, even in the case where agreement among the various predictions is relatively good (e.g., NB), there are differences of ca. 6–9 residues at each end, corresponding to ca. 2 turns of helix. So, the stability of different lengths of each TM helix when embedded in the middle of a bilayer was explored by multiple MD simulations. For each initial TM helix model, a sequence of 20 residues was chosen that matched as closely as possible to the consensus prediction (NB: IIITICVSLI^{30}VILIVFGCIA^{40}; CM2: L^{60}GLGIITMLYL^{70}LVKIIIELV; Vpu:VA^{10}LVVAIIIAIV^{20}VWSIVIIE). These models were then extended by approximately one or two α-helical turns at each end giving 28mer (NB: IRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFI; CM2: TLASL^{60}GLGIITMLYL^{70}LVKIIIELVN^{80}GFV; Vpu: IVAIVA^{10}LVVAIIIAIV^{20}VWSIVIIEYR^{30}KI) and 36mer models (NB: PITHIRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFINKNN; Vpu: QPIPIVAIVA^{10}LVVAIIIAIV^{20}VWSIVIIEYR^{30}KILRQR), respectively. For CM2, a 39mer (CM2: GYMLTLASL^{60}GLGIITMLYL^{70}LVKIIIELVN^{80}GFVLGRWERW^{90}) was generated to include the complete range of residues implicated in TM helix formation by sequence-based predictions.

Simulations and Secondary Structures

In all simulations, the structures deviate from the initial (i.e., in vacuo generated) models, as can be seen by the rise of the Cα RMSDs to ca. 0.1–0.2 nm within the first quarter of the simulation. This is typical for MD simulations of membrane proteins. The Cα-RMSD then remains constant for the remainder of the simulation, indicating overall helix stability on a nanosecond time scale. Similar behavior has been seen for simulations of TM helices formed by alamethicin[46] or by M2[36] or of the TM helices of bacteriorhodopsin.[48]

The time-dependent secondary structures (as defined by DSSP) of the TM domain models are shown in Fig. 2. Corresponding Cα traces, showing structures captured every 100 ps during the various simulations, are given in Fig. 3. For the NB 20mer (IIITICVSLI^{30}VILIVFGCIA^{40}) we find the α-helical conformation is conserved helix over the entire simulation (1 ns). Neither the bulky Phe-36 nor the hydrophilic Thr-24 and Ser-28 side-chains, all of which lie within the hydrophobic bilayer core, perturb the α-helicity of the model. The 28mer helix (IRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFI) also remains stable for the duration of simulation. However, a slight loss of helicity in favor of turn structures is observed around the first 3–4 residues including Ile-17 and Gly-19, during in the first half of the simulation. The time-averaged Φ and ψ backbone torsion angles (data not shown) deviate from canonical α-helical values for several of the resi-

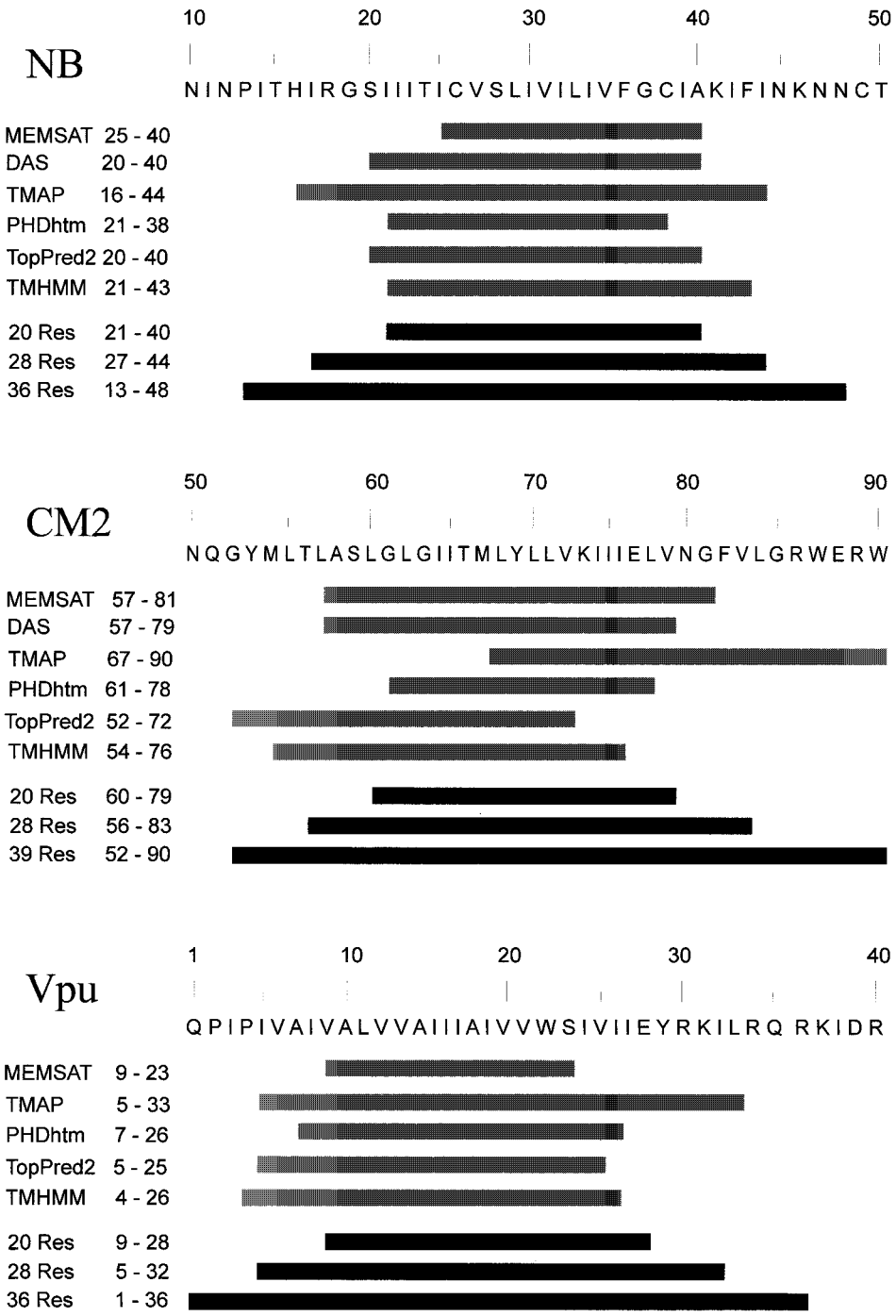

FIGURE 1 Predictions of TM helices using various algorithms (as indicated on the left-hand side of the figure). Numbers next to the names of each algorithm indicate the predicted helix start and helix end residues. Sequences of the regions of the three virus channel proteins around the predicted TM helices are shown at the top of each panel. Grey lines indicate the lengths of the predicted TM helices. Black lines indicate the three models generated for each protein, i.e., 20mer, 28mer, and 36mer (39mer for Vpu). For Vpu, DAS proposes a TM region from residue 30–53.

dues between Ile-17 and Ile-21. Since Ile-17 and Gly-19 are within the "interfacial" region defined by phospholipid headgroups (Fig. 4), these residues may be playing their traditional helix-destabilizing role when in a relatively polar environment[49] in this early stage of the simulation. There was no signif-

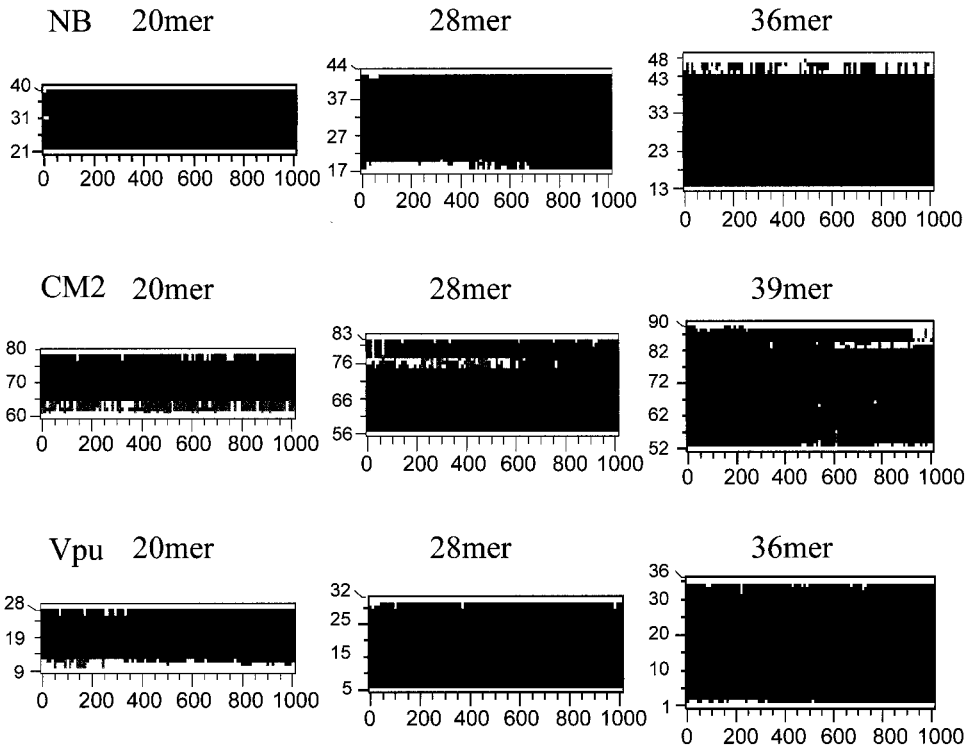

FIGURE 2 Analysis of secondary structure of the helices using DSSP algorithm.[56] For each model, secondary structure (vertical axis) is shown as a function of time (horizontal axis), using the following shading scheme: black = α-helix; dark grey = 3_{10}-helix; grey = turns; and white = random coil. The vertical axes list the residue numbers.

icant translational movement of the helix along the normal to the bilayer throughout the simulation. We suggest, therefore, that there is a localized interaction of these residues within the region of the polar lipid headgroups during the later part of the simulation, which then supports helix (re)formation. In the 36mer (PITHIRGS^{20}IIITICVSLI^{30}VILIVFGC IA^{40}KIFINKNN), we find mostly turns and some

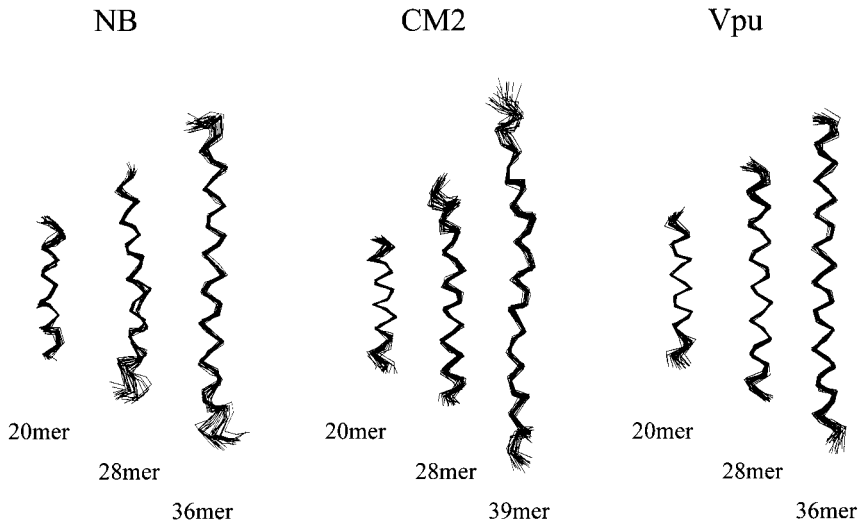

FIGURE 3 Superposition of Cα traces corresponding to structures saved every 100 ps. The structures are superimposed using their central cores, i.e., excluding 4 residues at each end.

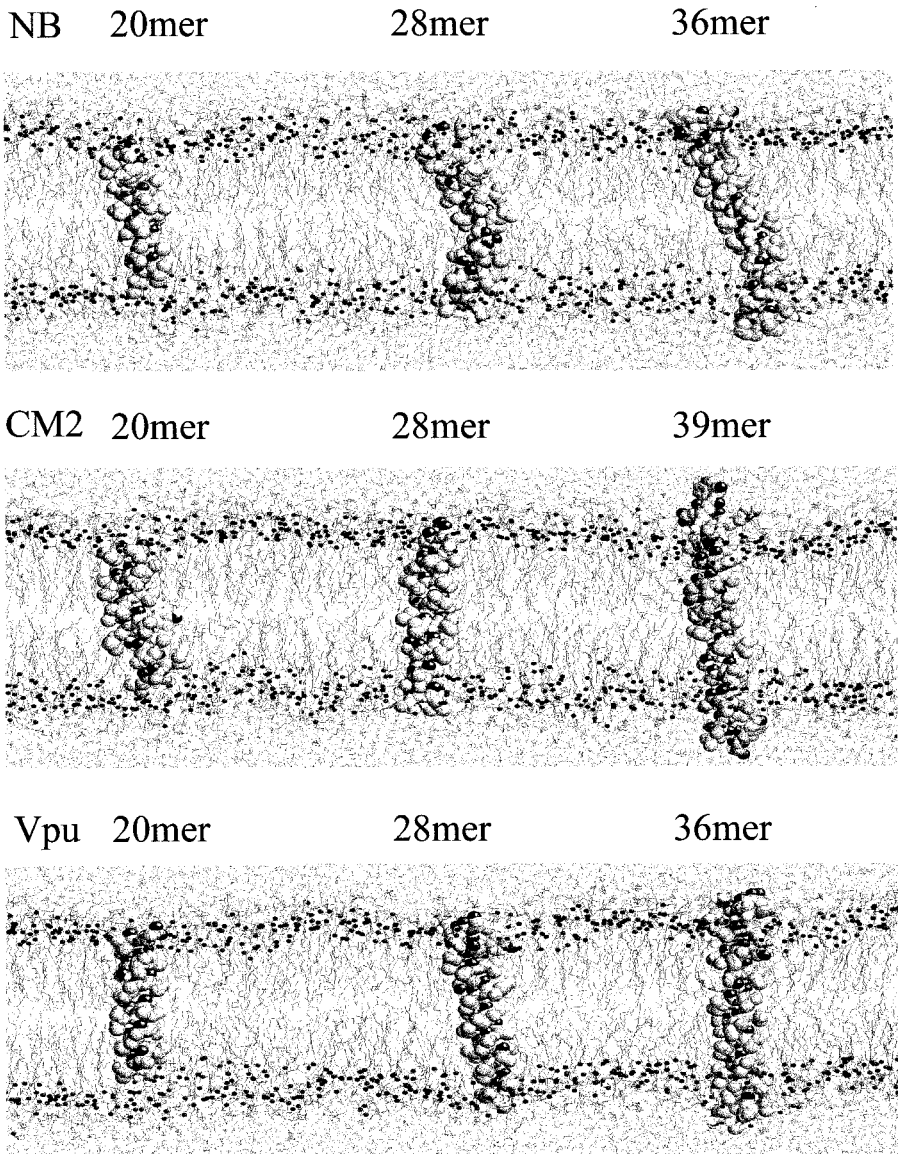

FIGURE 4 Images of the peptide/lipid/water systems at the ends of the simulations. The helices are shown in spacefilling format. The P atoms of the phospholipid head groups are shown as small spheres.

bends at the C terminus of the model. This break in the helix is due to hydrophilic residues (Asn-45, Lys-46, Asn-47, and Asn-48), which form hydrogen bonds with the phospholipid headgroups and interfacial water molecules. Their average Φ and ψ values deviate significantly from standard α-helical values.

The 20mer of CM2 (L^{60}GLGIITMLYL^{70}LV KIIIELV) is unstable in an α-helical conformation, adopting a mixture of mostly longer 3_{10}-helix and turn conformation at its N-terminus throughout the 1 ns simulation (Fig. 2). This may be related to the highly flexible residues Gly-61° (Φ-56°, ψ-37°), and Gly-63 (Φ-48°, ψ-25°) present in this region. These residues perturb the α-helical conformation within the lipid bilayer. In particular, in order to satisfy the hydrogen bonding of the terminus of the helix, there is partial unwinding at the glycines to increase its overall length. Extension of the CM2 sequence by 4 more residues at either end (giving the 28mer: TLA SL^{60}GLGIITMLYL^{70}LVKIIIELVN^{80}GFV) leads to a stable α-helix at the N-terminus (Fig. 2). However, in the C-terminal third of the molecule, there is a loss of α-helicity around residues 74–77 within the first 600 ps of the simulation. This perturbation disappears again towards the end of the 1-ns time period. For the

39mer (GYMLTLASL^{60}GLGIITMLYL^{70}LVKIIIE LVN^{80}GFVLGRWERW^{90}), the α-helical conformation is largely stable over the whole simulation time (Fig. 2), apart from at the C-terminus (residues Phe-82 to Arg-89), where there is a loss of α-helicity that starts at ca. 600 ps and continues for the remainder of the simulation.

The Vpu 20mer (VA^{10}LVVAIIIAIV^{20}VWSIVIIE) has a highly hydrophobic N-terminal half. The first 3–4 residues do not remain in a stable α-helical conformation within the bilayer. Instead, they adopt a turn conformation to span the bilayer. The conformation remains constant throughout the time course of the simulation. As was the case with CM2, the non-helical region contains a small side-chain, Ala-10, which adopts a more 3_{10}-like (Φ-57°, ψ-20°) conformation. The hydrophilic residues at the C-terminus do not perturb the α-helicity of the Vpu 20mer, even though the helix is buried within the bilayer core. The extended helix (28mer, IVAIVA^{10}LVVAIIIAIV^{20} VWSIVIIEYR^{30}KI) retains an almost completely α-helical structure throughout the simulation. Further extension of the sequence (to give a 36mer, QPIPIVAIVA^{10}LVV AIIIAIV^{20}VWSIVIIEYR^{30} KILRQR) does not lead to a substantial change in helix conformation throughout the duration of the simulation, although there is some transient loss of helicity at the N- and C-termini. The additional residues at the C-terminus of the helix are hydrophilic (Arg-34, Gln-35, and Arg-36). Since these residues are located in the aqueous environment, they may help to stabilize the local α-helical conformation via hydrogen bonding with the phospholipid headgroups or the water molecules.

Helix Distortions and Orientations

None of the transmembrane helices remains completely undistorted throughout the time course of its simulation. The 20mer NB helix (IIITICVSLI^{30} VILIVFGCIA^{40}) starts slightly bent at residue Val-31 with a helix kink angle of ca. 15° at the beginning of the simulation. This small kink remains throughout the 1-ns simulation (see Fig. 4). A slight tilt of the helix is also observed. A more pronounced kink is observed in the NB 28mer (IRGS^{20}IIITICVSLI^{30}VI LIVFGCIA^{40}KIFI). Domain motion analysis[50] reveals that residues Val-27 and Ser-28 form a molecular "hinge." This leads to a rotation of the C-terminal segment away from the bilayer normal [Fig. 4(A II)]. Although the initial model of the NB 36mer (PITHIRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFIN KNN) is an almost completely undistorted helix (kink angle ca. 5°), it becomes consistently kinked during the simulation, with a kink angle of ca. 30° at the end of the simulation. This reflects rotation of a compact segment within the bilayer (from residues Cys-26 to Val-35) around two smaller hinge segments on each side, giving the TM helix a twisted shape. The hinges are formed by residues 21–24 toward the N-terminus and residues 36–39 toward the C-terminus.

The helices of the CM2 20mer (L^{60}GLGIIT MLYL^{70}LVKIIIELV) and 28mer (TLASL^{60}GLGIIT MLYL^{70}LVKIIIELVN^{80}GFV) also exhibit limited curvature (kink angle ca. 10°) at the start of the simulation, which remains throughout the 1 ns [Fig. 4(B I and II)]. The 39mer (GYMLTLASL^{60}GL GIITMLYL^{70}LVKIIIELVN^{80}GFVLGRWERW^{90}) shows more complex behavior, with a maximum kink of the helix at around 700 ns, which then substantially disappears by the end of the simulation. This bend is due to movement of helical segments about two "hinges" from residues Ala-58 to Ile-64, and from Leu-78 to Val-79. In contrast to the (albeit limited) flexibility observed for the NB and CM2 helices, the three Vpu helix models remain unkinked and untilted relative to the bilayer normal throughout their respective simulations (Fig. 4).

Hydrogen Bonding Interactions of Side-Chains

All the TM segments studied have a hydrophobic core flanked by hydrophilic residues at either end. The presumed role of these amino acids is to interact within the interfacial (water + lipid headgroup) environment and thus stabilize the TM α-helix. Figure 5 shows the fraction of hydrogen bonding of the individual side-chains with the lipid headgroups, water, other amino acid side-chains, and peptide backbone.

The hydrophilic residues of the 20mers of all helices do not show any significant hydrogen bonding and are, therefore, not shown. This suggests that these 20mer helices are not sufficiently long to fully span a POPC bilayer. A common picture for all of the longer TM helix models is that the hydrophilic residues at the protein/lipid/water interface have relatively low f_H values ($f_H < 0.4$) (Fig. 5). This can be explained by a high number of possible partners for hydrogen bonding (q), i.e., a high fluctuation in hydrogen bonding partners. Hydrophilic side-chains like Thr and Ser, which are buried within the hydrophobic environment, satisfy their hydrogen bonding with backbone carbonyls. If there are no alternative partners for hydrogen bonding, such residues remain hydrogen bonded with these groups throughout the simulation, yielding a value of f_H between 0.4–0.8. It is interesting to note that Thr-56 and Ser-59 in the NB 28mer

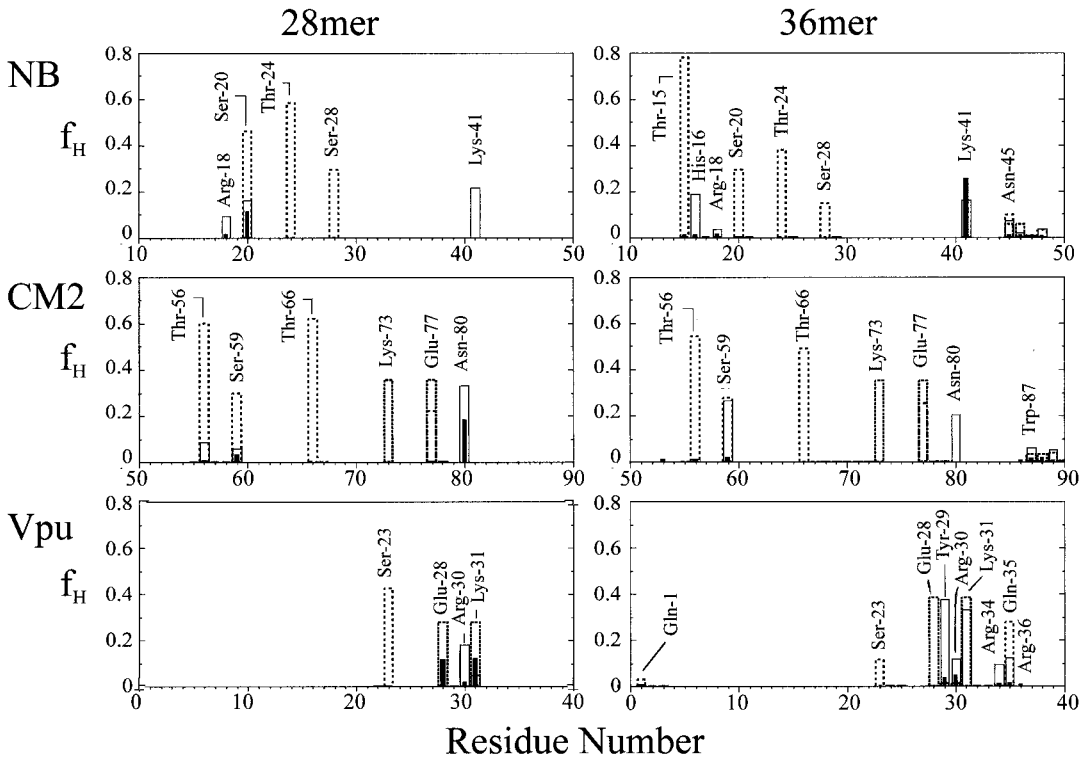

FIGURE 5 Fraction of hydrogen bonding (f_H) per residue for the side-chains of the 28mers and 36mers of NB, CM2, and Vpu. Open boxes indicate H-bonds to phospholipid headgroups; filled boxes indicate H-bonds to water molecules; long dotted-line boxes indicate H-bonds to the peptide backbone; and short dotted-line boxes indicate H-bonds to other side-chains. The fraction of hydrogen bonding is calculated as $f_H = n_H/(n_T \cdot q)$, where n_H is the total time during which a residue forms hydrogen bonds, n_T the total duration of the simulation, and q the number of observed numbers of hydrogen bonds formed by the residue. With a single type of hydrogen bond throughout the time course of the simulation, $f_H = 1$. A residue with hydrogen bonds to two different hydrogen bond acceptors or donors has a value of $f_H = 0.5$.

(IRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFI) and 36mer (PITHIRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}K IFINKNN) have a high fraction of hydrogen bonding to the backbone, despite the fact that they are close to the protein/lipid/water interface. This suggests that, at least within the timescale of these simulations, intramolecular H-bonding is favored over intermolecular H-bonding for these two residues. Ionized side-chains in NB and Vpu form salt bridges between their side-chains (NB: Lys-73 to Glu-77; Vpu: Glu-28 to Lys-31). In addition, Glu-77 (NB) and Lys-31 (Vpu) share H-bonds with the peptide backbone, the lipid headgroups, and water.

Stable TM Regions

From our data, we predict a stable α-helical conformation for the following transmembrane segments of the three proteins: (i) from residues 20–40 for NB (S^{20}IIITICVSLI^{30}VILIVFGCIA^{40}); (ii) from residues 56–76 for CM2 (TLASL^{60}GLGIITMLYL^{70}LVKIII); and (iii) from residue 5–25 for Vpu (IVAIVA^{10}L VVAIIIAIV^{20}VWSIV). Simulations reveal that the apparently unstable terminal residues of the 20mer peptides of NB (IIITICVSLI^{30}VILIVFGCIA^{40}) and CM2 (L^{60}GLGIITMLYL^{70}LVKIIIELV) retain their α-helical conformation, if they are embedded within a larger TM helix. Our simulation-based predictions are in good agreement with the consensus of the sequence-based TM predictions. In order to facilitate synthesis of putative TM segments for biophysical studies, the 28mers of all three proteins (NB: IRGS^{20}IIITICVSLI^{30}VILIVFGCIA^{40}KIFI; CM2: TLASL^{60}GLGIITMLYL^{70}LVKIIIELVN^{80}GFV; Vpu: IVAIVA^{10}LVVAIIIAIV^{20}VWSIVIIEYR^{30}KI) might be the peptide sequences of choice, as the

hydrophilic residues increase peptide solubility. Our simulations suggest that, in most cases, these TM segment "extensions" also adopt an α-helical conformation when in a lipid bilayer environment.

CONCLUSIONS

In this article, we investigated various models of TM helices derived from putative virus channel proteins. Neither the presence of glycine and β-branched residues in the TM segments nor the presence of occasional hydrophilic residues within a hydrophobic environment leads to any major loss of α-helicity. This is in agreement with the experimental studies of Deber et al.[51] and with principal components analysis of transmembrane segments, which stress the importance of β-branched side-chains.[52] The lipid surroundings stabilize the α-helical structures throughout the simulations. Breakdowns of helicity are mostly observed for hydrophobic residues in the vicinity of lipid headgroups. We conclude that, provided that the core region is hydrophobic, the exact sequence of amino acids within a predicted TM region may not affect helicity as much as the more polar residues at the flanks of this region.

Our results suggest that hydrophilic residues flanking both sides of a hydrophobic helical TM region anchor the helices, especially via H-bonds to the lipid headgroups. Mismatch of the spacing of the two groups of hydrophilic residues and the thickness of the bilayer would be expected to tilt and/or distort the helices.[53,54] The results of our study suggest that such a mismatch may sometimes lead to a change in peptide conformation and/or orientation, but that this is not inevitable. However, the abrupt termination of a helix within the lipid bilayer leads to a disturbance of the ends of this helix. This suggests that, in membrane protein modeling studies, it is essential to know the exact length of a TM helix. In general, our data suggest that it is better to have a longer helix segment rather then a shorter one. That is, helix/bilayer mismatch is tolerated if the helix is too long, but not if it is too short.[55]

Hydrophilic residues within the hydrophobic core of the bilayer (Ser, Thr, and Glu) saturate their hydrogen bonding capabilities either by forming hydrogen bonds with the side-chains of other hydrophilic residues or with the peptide backbone. This enables such residues to be accommodated within a bilayer, despite their hydrophilicity. However, such H-bond formation might lead to local distortion of the TM helix, as in the case of the CM2-28mer (Fig. 2), where there may be a correlation of H-bonding of the Glu-77 side-chain and the backbone with perturbation of the α-helix during the first half of the simulation. Such H-bonding with the peptide backbone might be modulated upon formation of bundles of helices in ion channel-forming assemblies.

For M2, there is strong evidence for a core helical TM region. Secondary structural analysis (Fig. 4 in Forrest et al., 1999[36]) indicates an optimal length of ca. 22 amino acids, which remains stably α-helical independent of the overall length of the TM segment modeled. Such clear evidence is not found for the TM helices of NB and CM2. Both the 28mer and the 36/39mers of these helices show deviation from helicity at their C-termini rather than at their N-termini. Loss of helicity in all segments, including that of M2, is found if more hydrophilic regions of the segment follow residues with β-branched side-chains (e.g., Ile) and extend into the lipid headgroup interface. Vpu seems to be an exception, in that there is little loss of helicity for its longer models. This reveals a possible trend. It seems that there may be a positive correlation between the content of β-branched (specifically, Ile/Val) residues and stability of α-helicity with a hydrophobic bilayer core. Thus, CM2 and M2 (both containing 7 Ile/Val residues in their cores) show greater deviations from α-helicity than does NB (11 Ile/Val) residues, which in turn is less stably α-helical than Vpu (15 Ile/Val residues). It occurs to us that this preliminary conclusion could be tested by assay (both experimentally and computationally) of suitably designed models of TM helices.

These simulations have been undertaken as a first step toward understanding the nature these putative viral ion channel proteins. They provide useful results on sequence features resulting in stable TM helices. Future studies will address the nature of helix/helix interactions within bundles, and how bundle structures are related to channel function.

This work was supported by a European Union (TMR) fellowship to WBF. MSPS thanks The Wellcome Trust for financial support. Our thanks to Indira Shrivastava and Rong-I Hong for helpful discussions. We thank the Oxford Supercomputing Centre for computer time.

REFERENCES

1. Sansom, M. S. P. Quart Rev Biophys 1993, 26, 365–421.
2. Tieleman, D. P.; Berendsen, H. J. C.; Sansom, M. S. P. Biophys J 1999, 76, 1757–1769.
3. Unwin, N. Nature 1995, 373, 37–43.
4. Doyle, D. A.; Cabral, J. M.; Pfützner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Cahit, B. T.; MacKinnon, R. Science 1998, 280, 69–77.

5. Chang, G.; Spencer, R. H.; Lee, A. T.; Barclay, M. T.; Rees, D. C. Science 1998, 282, 2220–2226.
6. Popot, J. L.; Engelman, D. M. Biochem 1990, 29, 4031–4037.
7. Lamb, R. A.; Pinto, L. H. Virol 1997, 229, 1–24.
8. Sansom, M. S. P.; Forrest, L. R.; Bull, R. Bioessays 1998, 20, 992–1000.
9. Kovacs, F. A.; Song, Z.; Wang, J.; Cross, T. A. Biophys J 1999, 76, A126.
10. Kukol, A.; Adams, P. D.; Rice, L. M.; Brünger, A. T.; Arkin, I. T. J Mol Biol 1999, 286, 951–962.
11. Sansom, M. S. P. Curr Opin Struct Biol 1998, 8, 237–244.
12. Duff, K. C.; Ashley, R. H. Virol 1992, 190, 485–489.
13. Duff, K. C.; Kelly, S. M.; Price, N. C.; Bradshaw, J. P. FEBS Lett 1992, 311, 256–258.
14. Duff, K. C.; Gilchrist, P. J.; Saxena, A. M.; Bradshaw, J. P. Virol 1994, 202, 287–293.
15. Sansom, M. S. P. Prog Biophys Mol Biol 1991, 55, 139–236.
16. Pinto, L. H.; Holsinger, L. J.; Lamb, R. A. Cell 1992, 69, 517–528.
17. Chizhmakov, I. V.; Geraghty, F. M.; Ogden, D. C.; Hayhurst, A.; Antoniou, M.; Hay, A. J. J Physiol 1996, 494, 329–336.
18. Shimbo, K.; Brassard, D. L.; Lamb, R. A.; Pinto, L. H. Biophys J 1996, 70, 1335–1346.
19. Hay, A. J.; Wolstenholme, A. J.; Skehel, J. J.; Smith, M. H. EMBO J 1985, 4, 3021–3024.
20. Wang, C.; Takeuchi, K.; Pinto, L. H.; Lamb, R. A. J Virol 1993, 67, 5585–5594.
21. Brassard, D. L.; Leser, G. P.; Lamb, R. A. Virol 1996, 220, 350–360.
22. Pekosz, A.; Lamb, R. A. Virol 1997, 237, 439–451.
23. Hongo, S.; Sugawara, K.; Muraki, Y.; Matsuzaki, Y.; Takashita, E.; Kitame, F.; Nakamura, K. J Virol 1999, 73, 46–50.
24. Betakova, T.; Nermut, M.; Hay, A. J Gen Virol 1996, 77, 2689–2694.
25. Hongo, S.; Sugawara, K.; Muraki, Y.; Kitame, F.; Nakamura, K. J Virol 1997, 71, 2786–2792.
26. Sunstrom, N. A.; Prekumar, L. S.; Prekumar, A.; Ewart, G.; Cox, G. B.; Gage, P. W. J Memb Biol 1996, 150, 127–132.
27. Chizhmakov, I.; Ogden, D.; Betakova, T.; Phillips, A.; Hay, A. Biophys J 1998, 74, A319.
28. Maldarelli, F.; Chen, M. Y.; Willey, R. L.; Strebel, K. J Virol 1993, 67, 5056–5061.
29. Schubert, U.; Bour, S.; Ferrer–Montiel, A. V.; Montal, M.; Maldarelli, F.; Strebel, K. J Virol 1996, 70, 809–819.
30. Paul, M.; Mazumder, S.; Raja, N.; Jabbar, M. A. J Virol 1998, 72, 1270–1279.
31. Friborg, J.; Ladha, A.; Goettlinger, H.; Haseltine, W. A.; Cohen, E. A. J AIDS Hum Retroviruses 1995, 8, 10–22.
32. Schubert, U.; Strebel, K. J Virol 1994, 68, 2260–2271.
33. Ewart, G. D.; Sutherland, T.; Gage, P. W.; Cox, G. B. J Virol 1996, 70, 7108–7115.
34. Schubert, U.; Ferrer–Montiel, A. V.; Oblatt–Montal, M.; Henklein, P.; Strebel, K.; Montal, M. FEBS Lett 1996, 398, 12–18.
35. Coady, M. J.; Daniel, N. G.; Tiganos, E.; Allain, B.; Friborg, J.; Lapointe, J. Y.; Cohen, E. A. Virol 1998, 244, 39–49.
36. Forrest, L. R.; Tieleman, D. P.; Sansom, M. S. P. Biophys J 1999, 76, 1886–1896.
37. Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. J Mol Biol 1998, 280, 925–932.
38. Duan, Y.; Kollman, P. A. Science 1998, 282, 740–744.
39. Jones, D. T.; Taylor, W. R.; Thornton, J. M. Biochem 1994, 33, 3038–3049.
40. Cserzo, M.; Bernassau, J.–M.; Simon, I.; Maigret, B. J Mol Biol 1994, 243, 388–396.
41. Persson, B.; Argos, P. J Mol Biol 1994, 237, 182–192.
42. von Heijne, G. V. J Mol Biol 1992, 225, 487–494.
43. Rost, B.; Fariselli, P.; Casadio, R. Prot Sci 1996, 5, 1704–1718.
44. Brünger, A. T. X-PLOR Version 3.1. A System for X-ray Crystallography and NMR. (Yale Univ Press, CT), 1992.
45. Kerr, I. D.; Sankararamakrishnan, R.; Smart, O. S.; Sansom, M. S. P. Biophys J 1994, 67, 1501–1515.
46. Tieleman, D. P.; Sansom, M. S. P.; Berendsen, H. J. C. Biophys J 1999, 76, 40–49.
47. Kabsch, W.; Sander, C. Biopoly 1983, 22, 2577–2637.
48. Woolf, T. B. Biophys J 1996, 70, A377.
49. Li, S. C.; Deber, C. M. FEBS Lett 1992, 311, 217–220.
50. Hayward, S.; Berendsen, H. J. C. Proteins: Struc Func Genet 1998, 30, 144–154.
51. Deber, C. M.; Li, S.–C. Biopoly 1995, 37, 295–318.
52. Koshi, J. M.; Bruno, W. J. Proteins: Struct Func Genet 1999, 34, 333–340.
53. de Planque, M. R. R.; Greathouse, D. V.; Koeppe II, R. E.; Schaefer, H.; Marsh, D.; Killian, J. A. Biochem 1998, 37, 9333–9345.
54. Webb, R. J.; East, J. M.; Sharma, R. P.; Lee, A. J. Biochem 1998, 37, 673–679.
55. Mouritsen, O. G.; Bloom, M. Ann Rev Biophys Biomol Struct 1993, 22, 145–171.
56. Kabsch, W.; Sander, C. Biopoly 1983, 22, 2577–2637.