Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 23, Issue Number 5, (2006)

Full Length Vpu from HIV-1: Combining Molecular Dynamics Simulations with NMR Spectroscopy

http://www.jbsdonline.com

V. Lemaitre[1,3]
D. Willbold[4]
A. Watts[1]
W. B. Fischer[1,2,*]

[1]Biomembrane Structure Unit
Department of Biochemistry
Oxford University
South Parks Road
Oxford OX1 3QU, UK
[2]Bionanotechnology Interdisciplinary Research Collaboration
Clarendon Laboratory
Department of Physics
Oxford University, Parks Road
Oxford OX1 3PU, UK
[3]Nestec S.A.
BioAnalytical Science Department
Vers-Chez-Les-Blanc
CH-1000 Lausanne 26, Switzerland
[4]Institute of Physical Biology and BMFZ
Heinrich Heine University, Düsseldorf
and IBI-2, Research Centre Jülich
D-52425 Jülich, Germany

Abstract

Based on structures made available by solution NMR, molecular models of the protein Vpu from HIV-1 were built and refined by 6 ns MD simulations in a fully hydrated lipid bilayer. Vpu is an 81 amino acid type I integral membrane protein encoded by the human immunodeficiency virus type-1 (HIV-1) and closely related simian immunodeficiency viruses (SIVs). Its role is to amplify viral release. Upon phosphorylation, the cytoplasmic domain adopts a more compact shape with helices 2 and 3 becoming almost parallel to each other. A loss of helicity for several residues belonging to the helices adjacent to both ends of the loop region containing serines 53 and 57 is observed. A fourth helix, present in one of the NMR-based structures of the cytoplasmic domain and located near the C-terminus, is lost upon phosphorylation.

Key words: Vpu; HIV-1; Membrane protein; Molecular dynamics simulations; NMR spectroscopy.

Introduction

Structural models are important in the process of elucidating the mechanistic details of protein function. A common procedure to reduce the complexity of a problem is to initially consider small structural components of the protein under investigation. Such an approach has been used to derive molecular models of large globular proteins (1, 2), and also of membrane proteins [(2), *e.g.*, for K^+ channel reviewed in (3)]. For membrane proteins this is often the only route available to obtain a molecular description, since many of them are difficult to crystallize.

In an attempt to construct a model of a type I integral membrane protein, the transmembrane (TM) structure, for which the structural motif is known from, *e.g.*, solid state NMR spectroscopy (4), is combined with a structural model of the cytoplasmic domain, which was resolved by solution NMR spectroscopy (5, 6). Both models are then joined by superimposing overlapping common parts. In the next stage the stability and dynamics of such a hypothetical 'full length' model was assessed using molecular dynamics (MD) simulations, in the presence of a hydrated lipid bilayer. This protocol is applied in the present study to Vpu, an 81 amino acid type I integral membrane protein encoded by HIV-1 (7, 8) and some closely related SIVs (9-11).

Vpu is involved in the amplification of viral release (12). It is active in the endoplasmic reticulum where it initiates the ubiquitinin-mediated proteolysis of CD4, the viral receptor protein [(13-15), for a recent review see (16)]. In the vicinity of the plasma membrane it is found to enhance the release of budding virions form the cell surface (12). Full length Vpu protein and peptides, which represent only the TM part, reconstituted into lipid membranes show channel activity (4, 15, 17, 18). For both functions, distinct domains of the protein are responsible. In the ER, the cytoplasmic domain with its two phosphorylation sites (Ser-53 and Ser-57) is

[*]Phone: +44 – 1865 – 275776
Fax: +44 – 1865 - 275234
Email: wolfgang.fischer@bioch.ox.ac.uk

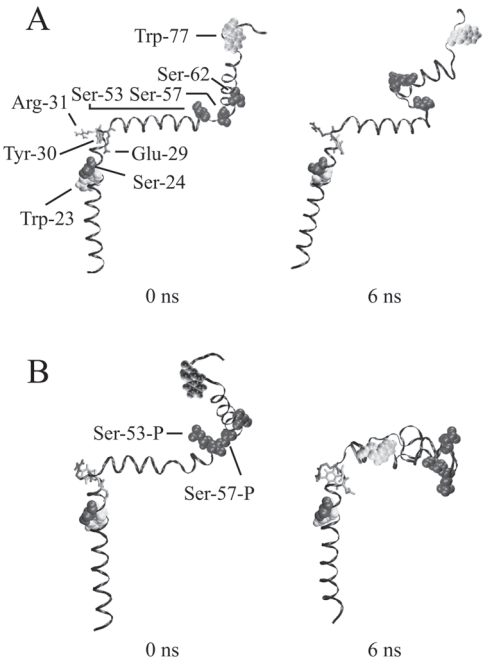

Figure 1: (**A**) Full length Vpu model, referred to as the HV1H2 model, based on the cytoplasmic structure from (23): the beginning (0 ns) and end (6 ns) of the simulations. All tryptophans (Trp-22, -76, light grey) and serines (Ser-23, -52, -56, -61, dark grey) are shown on their van der Waals representation, the EYR-motif [Glu-28 (light grey), Tyr-29 (grey), Arg-30 (dark grey)] are displayed in stick mode. Lipid and water molecules are omitted for clarity. (**B**) Phosphorylated model at the two serine sites 52 and 56 (lines). The shape and colour coding is the same as for (**A**).

required. Even though the phosphate groups are not essential for binding to CD4, they are required to initiate the degradation of CD4 (19). Ion channel activity is exclusively related to the TM domain of the protein (20).

There is a wealth of structural data for Vpu [reviewed in (5, 21)]. The structure of the cytoplasmic part, relevant for the present study, has been obtained from two independent studies on two related strains of HIV-1: HV1H2 (22, 23) and HV1S1 (22-24). A common feature of these studies is that the cytoplasmic part contains two helices, H-2 and H-3 (the N terminal TM helix is designated as H-1). These helices are connected by a flexible region containing the two phosphorylation sites Ser-53 and Ser-57. In one study (24), a fourth short helix at the immediate C terminal end is found. The two structural models derived from these investigations form the basis for our understanding of the cytoplasmic domain. These models are fairly consistent for the secondary structure, but differ in their tertiary structures.

The α-helical secondary structure of the TM domain has been determined experimentally and its orientation relative to the plane of the lipid bilayer has been estimated from both solid state NMR spectroscopy (4, 25) and FTIR spectroscopy (26). Solid state NMR spectroscopy was used to determine the approximate orientation of a second α-helix present in the cytoplasmic domain, with its long axis parallel to the membrane surface (27). Recent solution NMR spectroscopic data, recorded on a truncated Vpu peptide representing only the cytoplasmic part, reveal that upon phosphorylation H-3 is no longer detectable (28). The same authors have shown that a peptide, containing residues 50 to 57, adopts the same structural motif of a bend in this region independent of whether the data are recorded in solution or if the peptide is bound to the WD domain of β-TrCP (29).

In the present computational study explicit NMR spectroscopic data of the cytoplasmic domain of Vpu (23, 24) are used to generate models for full length Vpu. This work continues earlier MD simulations on a Vpu_{1-52} model (30).

These simulations give information (i) on the stability of the secondary structural elements of a putative full length Vpu protein when placed in a fully hydrated bilayer.

Materials and Methods

Hypothetical models for the full-length Vpu from HIV-1 were built for two different strains, HV1H2 and HV1S1:

Vpu HV1H2:
-QPIPIVAIV[10] ALVVAIIIAI[20] VVWSIVIIEY[30] RKILRQRKID[40] RLIDRLIERA[50] EDSGNESEGE[60] ISALVEMGVE[70] MGHHAPWDVD[80] DL

Vpu HV1S1:
AceQPLQILAIV[10] ALVVAAIIAI[20] VVWTIVYIEY[30] RKILRQRKID[40] RLIDRITERA[50] EDSGNESEGD[60] QEELSALVER[70] GHLAPWDVDD[80] L

The numbering scheme refers to the residue identification number for both strains. The initial conformation for the computational model of Vpu_{1-81} HV1H2 (Figure 1A,B) was obtained by superposition of the structure of Vpu_{32-81} HV1H2 (23) with the computational starting structure (0 ns) of the kinked model Vpu_{1-53} (30). The maximum possible overlap was not used because the residues at the termini of the peptide chain tend to lose their secondary structure, due to the increased mobility of those residues. Therefore, only the residues 41 to 48 were used to adjust and combine the two structures. In order to perform the adjustment, the distances between the backbone atoms of the two structures (CA, C, O, and N) was minimized using ProFit, and its least-square fitting algorithm. The structures were then subject to a short energy minimi-

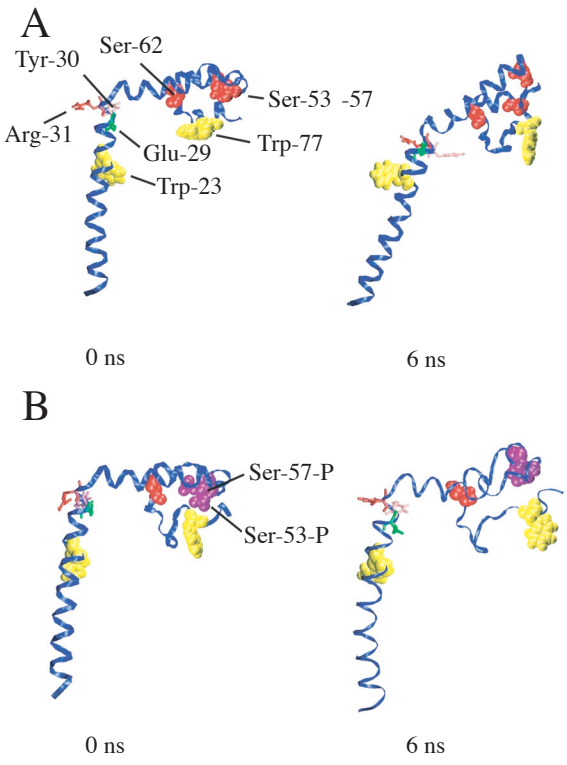

Figure 2: (**A**) the full length model HV1S1-sup (24) at the beginning (0 ns) and end (6 ns) of the simulations. (**B**) the full length phosphorylated model HV1S1-sup-P. The shape of the tryptophans, serines and the EYR-motif is the same as outlined in Figure 1. The color coding is red for the serines (and phosphorylated serines), yellow for the tryptophans, red (stick modus) for Arg-31, green for Glu-29, and light pink for Tyr-30.

zation in vacuum. The models are labelled as, *e.g.*, HV1H2 and HV1H2-P, if serines 53 and 57 are phosphorylated.

For the computational structure of Vpu_{1-81} HV1S1 (Figure 2A,B) the same procedure was applied by superimposing the Vpu_{37-81} experimentally derived structure (24) with the starting structure of the computationally derived Vpu_{1-52} structure (HV1H2) described in (30). The experimental structure Vpu_{37-81} was taken as the listed number one model out of the nine structures deposited at the RCSB Protein Data Bank [http://pdb.ccdc.cam.ac.uk/pdb/, entry name 1VPU (24)]. Swiss PDBviewer was used to build the HV1S1 TM section from the kinked HV1H2 Vpu_{1-52} structure (30). Point mutations were made so that the sequence of the peptide corresponded to the HV1S1 strain in this section. The structures were designated HV1S1-sup and HV1S1-sup-P depending on the presence of the phosphate

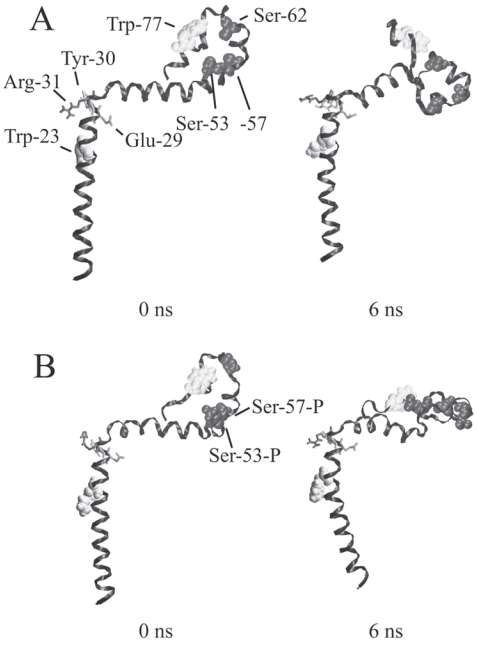

Figure 3: (**A**) the full length model HV1S1-rot (24) at the beginning (0 ns) and end (6 ns) of the simulations. (**B**) the full length phosphorylated model HV1S1-rot-P. The shape and color coding of the tryptophans, serines, and the EYR-motif is the same as outlined in Figure 1.

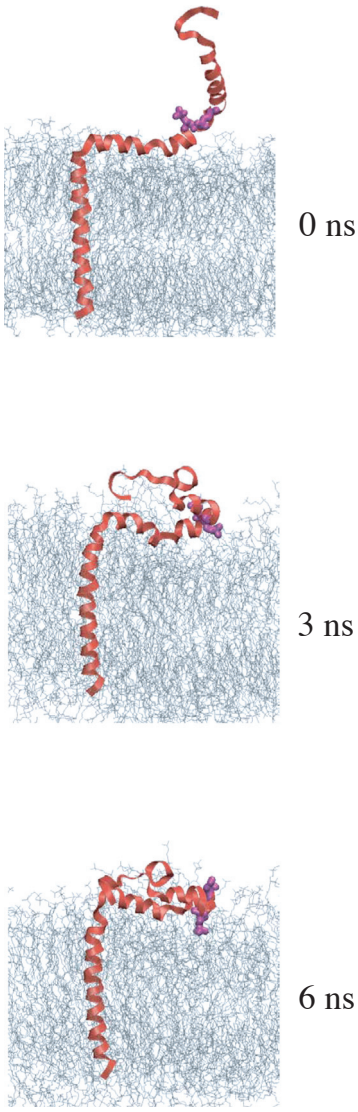

Figure 4: A section of the lipid bilayer (grey lines) around the HV1H2-P model (ribbon representation in red) at the beginning (0 ns) of the simulation and after 3 and 6 ns. Water molecules are omitted. The phosphorylated serines are shown as van der Waals radius spheres (purple).

groups at serines 53 and 57. As the superposition leads to a model in which the cytoplasmic part points towards the lipid bilayer, a second hypothetical structure (HV1S1-rot and HV1S1-rot-P, depending on phosphorylation state), was generated in which the cytoplasmic part was rotated around the peptide bond between residues 51-52 to bring the cytoplasmic part into the aqueous environment (Figure 2A,B). The structures for Vpu HV1S1 were then energy minimized.

For each strain, a structure for the protein with phosphorylated serine residues (residues 53 and 57) was built. First, the coordinates of a phosphorylated serine residue were built using PRODRG (31). Then, the phosphorylated serine residues 53 and 57 were replaced in the structures. This was achieved by fitting the phosphorylated residues on the corresponding serine residues of the models using ProFit as mentioned above (fit using CA, C, O, and N atoms). Finally, the coordinates of the fitted phosphate groups were added to the coordinate file of the structure and the models were energy minimized.

The phosphorylated serine residues are modelled using the force field fG43a1p, a modified fG43a1 force field (contributed by Dr. Graham Smith and available from the GROMACS website, www.gromacs.org/topologies/force_fields.php). The ionization state of each of the phosphate residues was chosen to be fully charged.

A large lipid bilayer consisting of 512 POPC molecules (1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phosphatidyl-choline), was created by adding 4 patches of a bilayer consisting of 128 POPC molecules [popc128a.pdb (http://moose.bio.ucalgary.ca/Downloads (32-34)] together. This procedure was followed by energy minimization of the hydrated lipid patch and a 100 ps equilibration run with the phosphorous atoms initially restrained to their z-position (1000 kJ/nm/mol) and gradually released. Finally, a 1 ns MD simulation was applied. The hole for the protein was created by overlaying the Vpu models and the bilayer. Lipid molecules whose phosphorous atoms overlapped with the GRASP representation of the peptide (Van der Waals surface) were removed.

For Vpu HV1H2, there are 238 lipids in the leaflet with the cytoplasmic part (242 for the phosphorylated Vpu HV1H2) and 246 lipids (247 for the phosphorylated Vpu HV1H2) in the other leaflet (484 lipids in total, 489 for the phosphorylated Vpu HV1H2). After insertion of Vpu_{1-81} HV1H2, the protein/lipid system (Figure 2A) was hydrated with approximately 69 water molecules per lipid independent of the phosphorylation state of the computational structure. The final overall number of atoms in the simulation is 126,256 (128,122 for the phosphorylated Vpu HV1-H2), including 6 Na^+ (10 for the phosphorylated Vpu) to neutralize the electric charges in the simulation box. The treatment of long range electrostatics requires an electrically neutral box.

For Vpu HV1S1, this leads to a removal of an unequal number of lipids on both sides of the bilayer, irrespective of the phosphorylation state, resulting in 238 lipids in the leaflet with the cytoplasmic part and 245 lipids in the other leaflet (483 lipids in total). After insertion of Vpu_{1-81} HV1S1, the protein/lipid system (Figure 2A) was hydrated with approximately 61 water molecules per lipid molecule. The overall number of atoms in the simulation at this stage was 114,143 (114,363 for phosphorylated Vpu HV1S1), including 6 Na^+ (10 for phosphorylated Vpu) to neutralize the electric charges in the simulation box. The large number of water molecules is necessary to include the bulky cytoplasmic part of this model and to avoid mirroring effects (the protein reflecting its image in the neighboring boxes), especially in the use of the particle-mesh Ewald (PME) method for including long range electrostatic interactions.

A 100 ps MD simulation with position-restraints on the backbone of the protein only was performed to equilibrate the system at 300 K and to allow the lipid mol-

ecules next to the cytoplasmic part of the protein to fill the space beneath it. The production phase was run at 300 K for 6 ns.

Gromacs 3.0.5 (35), including the simple point charge (SPC) water model (36), was used to perform the MD simulations. The simulations used a time step of 2 fs and a LINCS algorithm to keep the geometry of the molecules. An isothermal-isobaric ensemble (NPT) was used with periodic boundaries and Berendsen temperature and pressure coupling (p = 1 bar, T = 300K). Long-range electrostatic interactions have been calculated using the particle-mesh Ewald (PME) method. Lennard-Jones and short-range Coulombic interactions were cut off at 1.0 and 0.9 nm, respectively. The bilayer is placed in the center of a box with dimensions of 11.5 × 11.5 × 12.6 nm. This allows for a water layer with an average thickness of 4.5 nm on each side of the bilayer.

The secondary structure analysis is based on the program DSSP (define secondary structure of proteins, www.cmbi.kun.nl/gv/dssp/) which uses the presence or absence of hydrogen bonding as the decision parameter for defining the structural element (37).

The simulations were run on a Beowulf cluster, using four CPUs. Each node has Dual AMD Athlon 1.2GHz processors with 1GB memory/node. The fast interconnect *wulfkit* (www.wulfkit.com) uses a Dolphin SCI 2D torus board along with software provided by Scali (www.scali.com). The operating system of the cluster is Linux RedHat 7.1 with kernel 2.4 and MPI version 1.2 provided by Scali. The integrator was designed by WorkstationsUK (www.workstationsuk.co.uk) as a result of collaboration between AMD and the CCLRC-RAL e-Science Centre. VMD (www.ks.uiuc.edu/Research/vmd/) and Raster3d (www.bmsc.washington.edu/raster3d/html/raster3d.html) were used for the graphical representations of the structures.

Results

Investigations of the cytoplasmic part of Vpu using solution NMR spectroscopy have been performed with two different Vpu strains, HV1H2 (22, 23) and HV1S1 (24), which differ by three residues at the N terminus and a few residues, between positions 74 and 79 on the C terminus. In the NMR experiments the cytoplasmic parts were either dissolved in aqueous 50% tetrafluorethanol solution for the HV1H2 strain (22, 23) or in a solution with high salt concentration for the HV1S1 strain (24). Both models are consistent regarding the secondary structure of H-2 and H-3. While in the HV1H2 structure the C terminus adopts a turn motif, the HV1S1 structure adopts a short helix (H-4). The HV1S1 structure is also more spherical due to the screening of the charged residues caused by the high salt concentration. HV1H2 structure adopts a rod-like shape.

Vpu HV1H2 adopts a structure in which the long axis of helix 2 (H-2) is parallel to the membrane surface, and the cytoplasmic domain beyond H-2 points away from the membrane surface (Figure 1A). At the end of the simulation, the TM helix tilts, H-2 remains parallel to the bilayer surface, and the C-terminus stays in the 'aqueous phase'. Phosphorylation (HV1H2-P) initiates a moderate kink upon the TM helix during the simulation and the cytoplasmic part adopts a more compact shape (Figure 1B).

In HV1S1-sup (Figure 2A), H-2 is positioned along the bilayer surface with the short helix 3 (H-3, residues Gly-54 – Glu-63) pointing into the phospholipid headgroup region. After 6 ns the TM helix adopts a tilt angle of 10.8 ± 4.3° and seems to move away from the bilayer. H-2 moves away from the lipid bilayer surface leaving the helix/turn motif (H/T-4) in contact with the membrane surface. In the phosphorylated state (HV1S1-sup-P, Figure 2B), H/T-4 unwinds after equilibra-

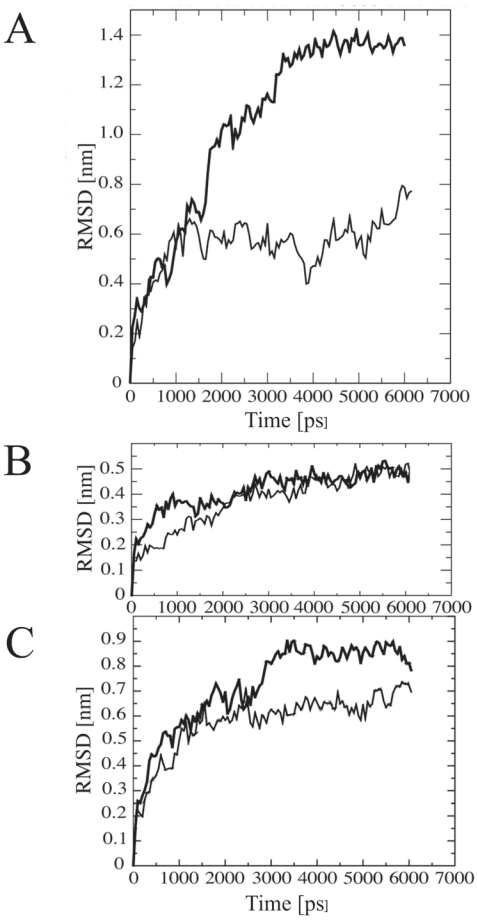

Figure 5: The root mean square deviation (RMSD) of the Cα atoms in respect to the starting structure for the non-phosphorylated (thin line) and phosphorylated (thick line) models. (**A**) HV1H2, (**B**) HV1S1-sup, (**C**) HV1S1-rot.

tion and remains unwound during the 6 ns production phase. At the end of the production phase, H-2 remains with its helix long axis parallel to the membrane surface. The residues towards the C-terminus unwind and remain in close proximity to the phospholipid head group region.

The cytoplasmic domain of the HV1S1-rot computational structure was rotated around Glu-51 and Asp-52 so that it points into the aqueous phase. The chosen bond lies between residues that are not within the H-2. It is assumed that this bond also allows for some flexibility *in vivo*. Similar to H1S1-sup, H-2 helix of HV1S1-rot moves so that at the end of the simulations the C terminus points away from the membrane surface (Figure 3A). Upon phosphorylation (HV1S1-rot-P), H-2 remains on the surface of the lipid membrane (Figure 3B) and, similar to HV1H2-P, the rest of the cytoplasmic remains in close proximity to the surface.

Snapshots of HV1H2-P embedded in the lipid bilayer at different time frames are shown in Figure 4. The lipid bilayer around HV1H2-P and around the other protein models remains fairly intact. However, the areas of the bilayer that are not in immediate contact with the protein experience some strong undulation so that the cytoplasmic part lies at the bottom of a well-like structure. This is found in all the models at the end of the simulations.

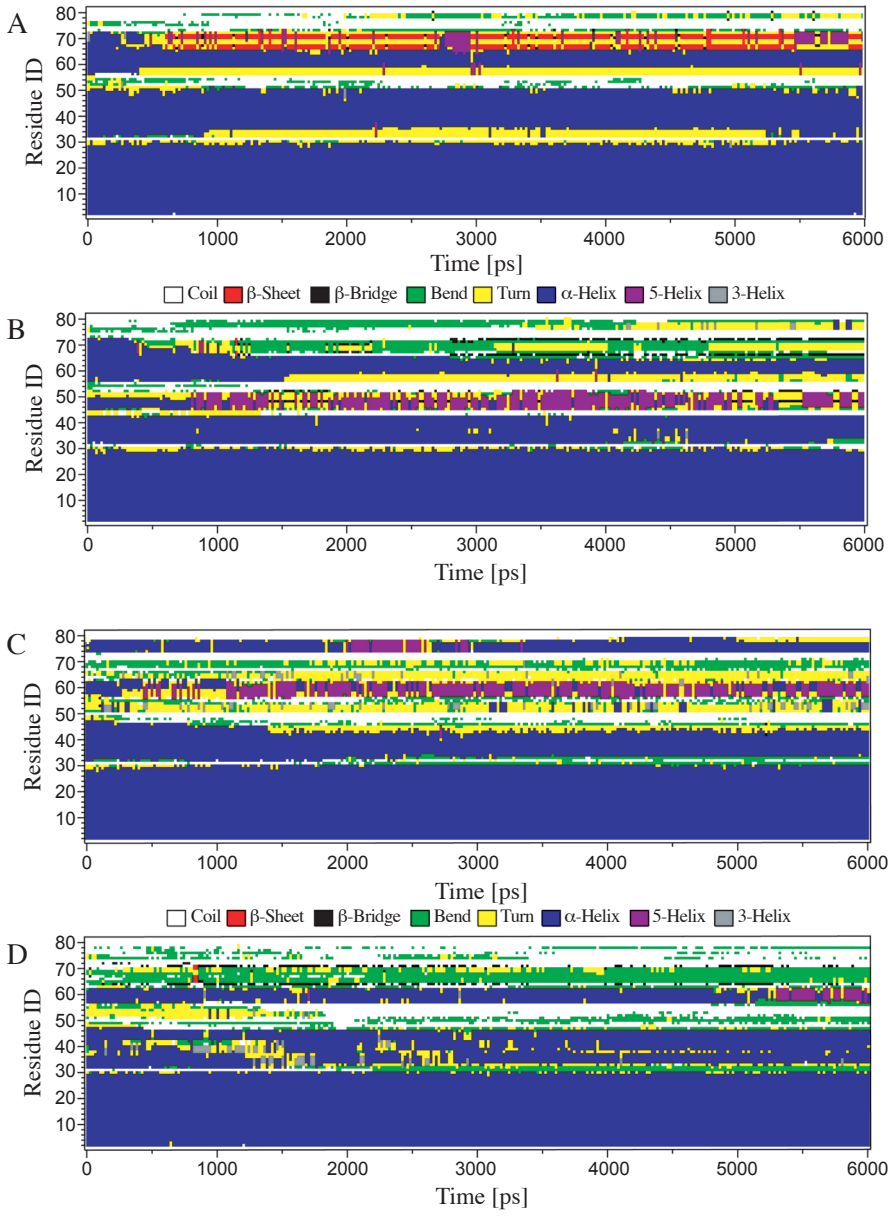

Figure 6: Secondary structural analysis of the protein models (**A**) HV1H2, (**B**) HV1H2-P, (**C**) HV1S1-rot, (**D**) HV1S1-rot-P, using DSSP (37). Color coding: White , coil; red, β-sheet; black, β-bridge; yellow, turn; green, bend; blue, α-helix; purple, 5-helix (π-helix), grey, 3_{10} helix.

The RMSD values level off in all simulations after approximately 3 ns (Figure 5). HV1H2 adopts similar values in the non-phosphorylated state of about 0.6 nm, while phosphorylation increases the value to 1.4 nm (Figure 5A). The high value is due to the high RMSD for H-3 relative to its own starting structure and also relative to the TM helix (data not shown). For the HV1S1-sup models, the level of RMSD values for the last 3 ns is around 0.5 nm (independently of the phosphorylation state, Figure 5B). For HV1S1-rot the levels are higher with approximately 0.6 nm and even approximately 0.85 nm for HV1S1-rot-P (Figure 5C). Similar to the values for HV1H2-P, the large movement of H-3 is responsible for the high values. As a consequence these structures strongly deviate from the initial starting structure.

The TM helix for all the non-phosphorylated models is tilted between 10º and 14º (14.3 ± 2.7º HV1H2, 10.8 ± 4.3º HV1S1-sup, 12.8 ± 3.6º HV1S1-rot, Table II). In the phosphorylated models, HV1H2-P and HV1S1-sup-P, the TM helices adopt an increased tilt angle of the TM helix (17.1 ± 4.7º and 19.9 ± 2.2º, respectively), whilst the tilt angle decreases for HV1S1-rot-P (7.8 ± 3.1º). For all the structures, phosphorylation induces a tilt between 20º and 30º of the C terminal end of H-2 into the bilayer away from its almost parallel initial alignment with the membrane surface. The kink angle of all the models remain in a range of 3.7º – 10º (Table II) which is also found by NMR spectroscopy of Vpu fragments including the TM segment.

In Vpu HV1H2, the complex salt bridge formed by Glu-29, Lys-32, and Arg-35 as described previously (30) (named Glu-28, Lys-31, and Arg-34 in (30)), is also found in this study. An average distance for salt bridges formed by Glu-29/Lys-32 and Glu-29/Arg-35 is calculated as 0.23 ± 0.01 nm and 0.23 ± 0.01 nm, respectively, with only modest change upon phosphorylation. For HV1S1-rot the salt bridges are less stable. In HV1S1-rot the two residues Glu-29 and Arg-35 retain the salt bridge between them with an averaged distance of 0.22 ± 0.03 nm. The salt bridge of Glu-29/Lys-32 breaks after about 2 ns, due to an increase in distance of more than 0.5 nm. In HV1S1-rot-P neither of the salt bridges exist (all average values are larger than 0.3 nm). After about 3 ns Glu-29 and Lys-38 form a salt bridge with values around 0.25 nm. The revolution of H-2, as mentioned for $Vpu_{1\text{-}52}$ (30), has been observed to a lesser extent in all the models. This might be caused by the larger cytoplasmic part used in the present models.

The high charge-density of the phosphate groups of the phosphorylated serines, each of -2, is compensated by two Na^+ per group continuously interacting with two of it. On average, up to two Na^+ are approaching the bilayer patch within a range of 0.22 nm during each of the simulation.

Evaluation of the Secondary Structure with DSSP

All models in common is a stable helical TM region ranging from Glu-2 to Tyr-30. In the HV1H2 structure, (Figure 6A, Table I) H-2 is defined from Lys-32 to Glu-51 and H-3 from Glu-56 to Gly-72. Within the 6 ns production phase two to three residues towards the N terminal ends of H-2 and H-3 unwind as well as part of the N terminal end of H-3. The remaining helical sections are listed in Table I. H-4 is not defined in this model. Phosphorylation (HV1H2-P, Figure 6B, Table I) leaves the TM helix unchanged (from Gln-2 – Tyr-30) at the end of the simulation. H-2 remains stable from Ile-32 to Ala-50. H-3 reaches at the end of the simulation from Gly-59 to Glu-66 as compared to the beginning of the simulation when it stretches from Glu-56 to Gly-72.

The analysis of secondary structure as a function of time for HV1S1-sup reveals the presence of four α-helices (Table I). The first helix, which is the membrane-spanning helix, is defined from Gln-2 to Val-26 from 0.5 ns until 5.5 ns. Helix-2 is fairly well defined from residues Ile-33 to Arg-49 throughout the simulation. H-3, which is defined from Asp-58 to Glu-62 at the beginning of the simulation, grows to include Asn-55 to Glu-62 for the last 3 ns of the simulation. H-4 includes Ala-74 to

Table I

The length of the individual helices in Vpu derived from DSSP analysis (37): The transmembrane (TM) helix, referred to as helix-1 (H-1), helix-2 (H-2), which is assumed to be parallel to the membrane surface, helix-3 (H-3), and a short helix-4 (H-4) at the C terminus. P denotes the respective phosphorylated state of the respective models. The data from this study are compared with those found in the literature.

	TM helix (H-1)	H-2	H-3	H-4 / Turn	References * this work
HV1H2					
HV1H2	Gln-2 – Tyr-30	Ile-33 – Glu-51	Glu-60 – Glu-66		*
HV1H2-P	Gln-2 – Tyr-30	Ile-33 – Arg-44	Gly-59 – Glu-66		*
HV1S1					
HV1S1-sup	Gln-2 – Val-26	Ile-33 – Arg-49	Asn-55 – Glu-62	Ala-74 – Asp-79	*
HV1S1-sup-P	Gln-2 – Glu-29	Lys-32 – Arg-45	Asp-60 – Leu-64	-	*
HV1S1-rot	Gln-2 – Tyr-30	Leu-34 – Asp-44	Ser-57 – Gln-62	Ala-74 – Val-78	*
HV1S1-rot-P	Gln-2 – Tyr-30	Leu-34 – Ile-46	Ser-57 – Gln-63	-	*
MD	Ile-5 – Val-25				(51)
MD	Gln-5 – Ile-26	Ile-32 – Glu-47	Glu-59 – Glu-69		(39)
NMR		Ile-42 – Glu-50	Glu-57 – Glu-69	Ala-74 – Asp-77 (T)	(22)
CD	Glu-5/Ile-6 – Tyr-29				(52)
NMR		Lys-37 – Asp-51	Glu-57 – His-72	His-73 – Val-78 (T)	(23)
NMR		Asp-40 – Ala-50	Asp-60 – Val-68	Pro-75 – Asp-79 (H)	(24)
SSNMR/NMR	Ala-9 – Tyr-29	Arg-30 – Ala-49	Gly-58 – Leu-70		(4)
SSNMR	Glu-5 – Glu-28				(25)

Table II

Tilt, kink, and crossing angles derived from the average values taken from 0 ns to 6 ns. For HV1H2, model-1 is the working model. For the model HV1S1, -sup stands for the superimposed model, where the cytoplasmic part is superimposed with the kinked model from (30) and -rot stands for the cytoplasmic part being rotated so that it faces the aqueous solution. P denotes the respective phosphorylated state of the respective models. The crossing angle is defined as the angle between the TM helix and helix H-2.

	Tilt		Kink		
	TM helix [°]	H-2 [°]	TM helix [°]	H-2 [°]	Crossing angle [°]
HV1H2					
Model-1	14.3 ± 2.7	91.3 ± 6.8	6.4 ± 2.4	5.4 ± 2.3	42.2 ± 4.2
P	17.1 ± 4.7	120.8 ± 9.2	7.4 ± 3.1	5.6 ± 2.2	21.4 ± 4.2
HV1S1					
-sup	10.8 ± 4.3	60.9 ± 5.8	3.7 ± 1.8	5.5 ± 2.6	54.3 ± 5.5
P	19.9 ± 2.2	83.9 ± 3.6	8.4 ± 2.3	19.4 ± 6.6	44.4 ± 2.9
-rot	12.8 ± 3.6	81.4 ± 6.9	5.7 ± 2.3	4.3 ± 3.0	46.9 ± 6.2
P	7.8 ± 3.1	108.7 ± 7.8	10.1 ± 2.5	23.6 ± 4.3	22.6 ± 4.8

Asp-79 throughout the simulation. Upon phosphorylation at the end of the simulation the TM helix also includes residues Gln-2 and Glu-29 in the model HV1S1-sup-P. However, helical motifs are now found between residues Lys-32 to Arg-45 (H-2) and Asp-60 to Leu-64 (H-3). The fourth helix is no longer detected.

At the end of the simulation the TM helix of HV1S1-rot (Figure 6C and Table I) H-2 stretches from Leu-34 to Asp-44 which comprises a shortening at its C-terminal end by three residues (Arg-45, Leu-46, Ile-47). H-3 is left as a short π-helix from Ser-57 to Glu-62, while H-4 includes residues Ala-74 to Val-78. Phosphorylation in HV1S1-rot-P (Figure 5D, Table I) seems to confirm the regions for H-2 (Leu-34 to Ile-46), and H-3 (Ser-57 to Glu-63). In HV1S1-sup-P, a total unwinding of H-4 is observed. The preferred structural motif in this region for HV1S1-sup-P and HV1S1-rot-P is a coiled structure with a minor turn motif. In this model a decrease of the helices at both sides of the two phosphorylation sites is not observed.

In summary, there is evidence that phosphorylation reduces the extent of helicity in H-2 and H-3 at their extremities adjacent to the two phosphorylation sites. H-4 defined in HV1S1 is not stable if the protein is phosphorylated.

Assessment of the Computational Models

The current computational structures incorporate all the structural information available to date based on experimental findings. Nevertheless, since a full length structure as such has not yet emerged, some compromise was necessary to finally model the full length Vpu. The merger of the cytoplasmic parts with a computational kinked structure is in this respect the most promising solution. The model HV1S1-sup is chosen as a 'test case' to monitor the events of a possible, but unlikely, orientation of the cytoplasmic domain. In that orientation the cytoplasmic domain moves out of the bilayer within the 6 ns simulation, indicative for this energetically unfavorable orientation. Furthermore, the force field used for the simulations is gromos43A2 which is improved for simulations of lipids (38). The secondary structural analysis regarding the stretch of the helices was based on DSSP (37), which evaluates the structure according to its hydrogen bond pattern. Keeping the same simulation parameters constant allows a direct comparison of the different models of full length Vpu. The model HV1S1-rot comprises an artificial intervention in the full length model. The bond around which the rotation was chosen to be done is plausible, however, it reflects only one conformation out of many which may be adopted by the protein *in vivo*. Consequently the current choice may bias the full length HV1S1 model into a narrow range of conformational space.

Due to the enormous CPU time necessary per single model, multiple runs, as suggested previously to evaluate simulation data (30), have not been done.

In the simulations, H-2 is embedded within the lipid headgroup region in order to account for the amphiphilic character of this helix, while H-3 is exposed to the aqueous phase. In the absence of the phosphate groups on Ser-53 and Ser-57, H-3 remains mostly in the aqueous phase. In the models with the phosphate groups H-3 and the rest of the cytoplasmic domain, a more compact fold is adopted which aligns H-3 almost parallel to H-2. A similar behavior during MD simulations on the cytoplasmic domain has been reported for a model which lacks the phosphate groups (39). However, the starting structure for that simulation was different and the cytoplasmic domain was initially positioned on top of the bilayer. Phosphorylation of the serine residues shortens H-2 towards its C terminus. The complex salt bridge formed between Glu-29, Lys-32, and Arg-35 in the kinked region between the TM helix and H-2 at the beginning of the simulation is broken upon phosphorylation. In conclusion, phosphorylation of the serines affects the secondary structure of Vpu. Strong structural arrangement of proteins in the presence of phosphate groups has also been reported for guanine binding proteins [reviewed in (40)].

The overall tilt angles for the TM helix of all models fall within the most recently-derived experimental value of around 15º (25), even though no tilt was imposed in the starting structures. This suggests that the tilt angle for the TM helix is independent of the rest of the molecule and possibly dependent on the lipid composition of the bilayer instead. The observed kink of the TM helix in this simulation [see also (30, 41)] is also observed experimentally (25).

It is recognized that the unequal numbers of lipids on both sides of the leaflets impose some unequally-distributed lateral pressure on the membrane, a situation which has been discussed previously (30). The unequal number of lipids in the two sides of the lipid bilayer might bias the protein model in its behavior so that the hydrophobic part of H-2 moves towards the hydrophobic part of the bilayer. Given the size of the system chosen, we decided to accept the possibility of such a biased protein-lipid interaction. The data reveal that the protein is not forced into an unnatural or unusual position during the simulation.

To fully understand the dynamics of a membrane protein, the lipid environment must be considered (42-44), especially in the case of a small protein such as Vpu. The lipid bilayer in the present study retains its characteristic structure around the proteins as an assembly of two parallel leaflets with their hydrophobic sections facing each other. However, undulations at the surface of the bilayer, adopting an overall wave-like form, are observed. Such undulations have been reported for MD simulations (10 ns) on very large bilayer patches consisting of 1024 dipalmitoylphosphatidylcholine (DPPC) molecules (45). In that study (45) the undulation seems to be dependent on the content of cholesterol. The simulations reported here are run without any cholesterol and with smaller lipid patches. However, the patches seem to be large enough for this type of behavior to be observed. On the other hand, the presence of ions, *e.g.*, sodium chloride, has been found to impose a considerable local effect on lipid membranes using even smaller patches (*e.g.*, 128 lipid molecules) with respect to lipid mobility, ion penetration depth into the head group region, and bilayer thickness (46, 47). The presence of such a low concentration of ions, (10 ions per 33,500 water molecules), is unlikely to affect the properties of the bilayers used in the present simulation.

Comparison of the Protein Models with Previous Computational and Experimental Results

The hypothetical structures presented in this study have much in common with the recent models derived from experiments (18, 27) and other simulations (39). The cytoplasmic domain, or at least H-2, remains in close contact with the lipid surface (39). H-2 seems to be fairly stable, but phophorylation decreases the helical content towards the C terminus. This behavior is confirmed by solution NMR spectroscopy studies of the phosphorylated fragment Vpu_{41-62} (28, 48). The structure of this particular segment seems not to change if it is also bound to β-TrCP protein (29). H-2 is aligned with its long axis parallel to the membrane surface according to solid state NMR spectroscopic studies with uniformly ^{15}N labelled Vpu (27) as well as with ^{15}N labelled peptides of Vpu (^{15}N-Leu-45-Vpu_{27-57}) (49). Upon phosphorylation, this helix in Vpu_{27-57} (49) additional broad components in the NMR spectra were observed. These components indicate the presence of various other amide orientations of the peptide in this sample and could agree with the distortion of the helicity in H-2 suggested in the present computational study. The NMR spectra of a peptide including H-3, ^{15}N-Ala-62-Vpu_{51-81}, indicates various sets of orientations independent of the phosphorylated state, which suggest a loose contact between H-3 and the lipid bilayer. A second solid state NMR spectroscopic study concluded that the cytoplasmic helices are orientated with the helix long axis parallel to the membrane surface, both of them in contact with the lipid membrane (27). X-ray reflectivity data from Langmuir monolayers suggest a strong dependence of the orientation of the cytoplasmic part of a full length construct on the surface pressure (50). Low surface pressure would lead to a large tilt angle of the TM helix compared to the membrane headgroup surface, and a disconnection of the cytoplasmic part from the membrane surface. Increasing surface pressure brings the cytoplasmic part back to the membrane surface, therefore supporting the model proposed by solid state NMR spectroscopy (27). At larger surface pressure with H-3 on top of H-2 is consistent with the experimental results. Thus, whilst H-2 is in fairly stable contact with the membrane surface, H-3 might also be in membrane contact, but would be very mobile. The results obtained in a 6 ns simulation, as shown in the present study, support the findings that the orientations of all the cytoplasmic helices in the phosphorylated models are along the membrane surface.

Depending on the experimental or computational set up used, there is some variability in the computed structure of the ends of the secondary structural elements. Nevertheless, the extent of these elements is similar.

As a conclusion, Vpu consists predominantly of two helices. In two out of the three models presented here a loss of helicity for several residues belonging to the helices adjacent to both ends of the loop region containing serines 53 and 57 is observed. In respect of H-2, the models built from orientational restraints measured by NMR spectroscopy are robust and support the idea of H-2 as a peptide float (30).

Acknowledgement

We thank Victor Wray, Braunschweig, Germany, for providing us with experimentally derived structural data of the cytoplasmic domain of Vpu. V. L. thanks BBSRC for an Industrial CASE scholarship. W. B. F. and A. W. acknowledge the Bionanotechnology IRC, MRC and BBSRC for funding this work. We thank CCLRC-RAL for computer facilities and service. Thanks to Peter Judge (Oxford) for stimulating discussions.

Reference and Footnotes

1. I. D. Campbell and A. K. Downing. *Nature Struc. Biol. NMR Supplement*, 496-499 (1998).
2. I. D. Campbell. *Biochem. Soc. Trans. 31*, 1107-1114 (2003).
3. P. C. Biggin, T. Roosild, and S. Choe. *Curr. Opin. Struc. Biol. 10*, 456-461 (2000).
4. C. Ma, F. M. Marassi, D. H. Jones, S. K. Straus, S. Bour, K. Strebel, U. Schubert, M. Oblatt-Montal, M. Montal, and S. J. Opella. *Prot. Sci. 11*, 546-557 (2002).
5. W. B. Fischer. *FEBS Lett. 552*, 39-46 (2003).
6. W. B. Fischer. *Viral Membrane Proteins: Structure, Function and Drug Design, Vol. 1*. Kluwer Academic/Plenum Publisher, New York (2005).
7. K. Strebel, T. Klimkait, and M. A. Martin. *Science 241*, 1221-1223 (1988).
8. E. A. Cohen, E. F. Terwilliger, J. G. Sodroski, and W. A. Haseltine. *Nature 334*, 532-534 (1988).
9. T. Huet, R. Cheynier, A. Meyerhans, and G. Roelants. *Nature 345*, 356-359 (1990).
10. V. Courgnaud, M. Salemi, X. Pourrut, E. Mpoudi-Ngole, B. Abela, P. Auzel, F. Bibollet-Ruche, B. Hahn, A.-M. Vandamme, E. Delaporte, and M. Peeters. *J. Virol. 76*, 8298-8309 (2002).
11. V. Courgnaud, B. Abela, X. Pourrut, E. Mpoudi-Ngole, S. Loul, E. Delaporte, and M. Peeters. *J. Virol. 77*, 12523-12534 (2003).
12. T. Klimkait, K. Strebel, M. D. Hoggan, M. A. Martin, and J. M. Orenstein. *J. Virol. 64*, 621-629 (1990).
13. R. L. Willey, F. Maldarelli, M. A. Martin, and K. Strebel. *J. Virol. 66*, 7193-7200 (1992).
14. T. Kimura, M. Nishikawa, and A. Ohyama. *J. Biochem. 115*, 1010-1020 (1994).
15. U. Schubert, S. Bour, A. V. Ferrer-Montiel, M. Montal, F. Maldarelli, and K. Strebel. *J. Virol. 70*, 809-819 (1996).
16. K. Levesque, A. Finzi, J. Binette, and E. A. Cohen. *Curr. Hiv Res. 2*, 51-59 (2004).
17. G. D. Ewart, T. Sutherland, P. W. Gage, and G. B. Cox. *J. Virol. 70*, 7108-7115 (1996).
18. G. G. Kochendörfer, D. H. Jones, S. Lee, M. Oblatt-Montal, S. J. Opella, and M. Montal. *J. Am. Chem. Soc. 126*, 2439-2446 (2004).
19. S. Bour, U. Schubert, and K. Strebel. *J. Virol. 69*, 1510-1520 (1995).
20. U. Schubert, A. V. Ferrer-Montiel, M. Oblatt-Montal, P. Henklein, K. Strebel, and M. Montal. *FEBS Lett. 398*, 12-18 (1996).
21. M. Montal. *FEBS Lett. 552*, 47-53 (2003).
22. V. Wray, T. Federau, P. Henklein, S. Klabunde, O. Kunert, D. Schomburg, and U. Schubert. *Int. J. Peptide Protein Res. 45*, 35-43 (1995).
23. T. Federau, U. Schubert, J. Floßdorf, P. Henklein, D. Schomburg, and V. Wray. *Int. J. Peptide Protein Res. 47*, 297-310 (1996).
24. D. Willbold, S. Hoffmann, and P. Rösch. *Eur. J. Biochem. 245*, 581-588 (1997).
25. S. H. Park, A. A. Mrse, A. A. Nevzorov, M. F. Mesleh, M. Oblatt-Montal, M. Montal, and S. J. Opella. *J. Mol. Biol. 333*, 409-424 (2003).
26. A. Kukol and I. T. Arkin. *Biophys. J. 77*, 1594-1601 (1999).
27. F. M. Marassi, C. Ma, H. Gratkowski, S. K. Straus, K. Strebel, M. Oblatt-Montal, M. Montal, and S. J. Opella. *Proc. Natl. Acad. Sci. Usa 96*, 14336-14341 (1999).
28. G. Coadou, N. Evrard-Todeschi, J. Gharbi-Benarous, R. Benarous, and J.-P. Girault. *Int. J. Biol. Macromolecues 30*, 23-40 (2002).
29. G. Coadou, J. Gharbi-Benarous, S. Megy, G. Bertho, N. Evrard-Todeschi, E. Segeral, R. Benarous, and J.-P. Giraudat. *Biochemistry 42*, 14741-14751 (2003).
30. I. Sramala, V. Lemaitre, J. D. Faraldo-Gomez, S. Vincent, A. Watts, and W. B. Fischer. *Biophys. J. 84*, 3276-3284 (2003).
31. D. M. F. van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hooft, and G. Vriend. *J. Comput. Aided Mol. Des. 10*, 255-262 (1996).

32. D. P. Tieleman, H. J. C. Berendsen, and M. S. P. Sansom. *Biophys. J. 76*, 3186-3191 (1999).
33. D. P. Tieleman, H. J. C. Berendsen, and M. S. P. Sansom. *Biophys. J. 76*, 1757-1769 (1999).
34. D. P. Tieleman, M. S. P. Sansom, and H. J. C. Berendsen. *Biophys. J. 76*, 40-49 (1999).
35. E. Lindahl, B. Hess, and D. Van der Spoel. *J. Mol. Mod. 7*, 306-317 (2001).
36. H. J. C. Berendsen, J. R. Griegera, and T. P. Straatsma. *J. Phys. Chem. 91*, 6269-6271 (1987).
37. W. Kabsch and C. Sander. *Biopolymers 22*, 2577-2637 (1983).
38. O. Berger, O. Edholm, and F. Jähnig. *Biophys. J. 72*, 2002-2013 (1997).
39. F. Sun. *J. Mol. Model. 9*, 114-123 (2003).
40. I. R. Vetter and A. Wittinghofer. *Science 294*, 1299-1304 (2001).
41. F. S. Cordes, A. Tustian, M. S. P. Sansom, A. Watts, and W. B. Fischer. *Biochemistry 41*, 7359-7365 (2002).
42. H. Heller, M. Schaefer, and K. Schulten. *J. Phys. Chem. 97*, 8343-8360 (1993).
43. D. P. Tieleman and H. J. C. Berendsen. *J. Chem. Phys. 105*, 4871-4880 (1996).
44. E. Jakobsson. *Trends Biochem. Sci. 22*, 339-354 (1997).
45. C. Hofsäss, E. Lindahl, and O. Edholm. *Biophys. J. 84*, 2192-2206 (2003).
46. R. A. Böckmann, A. Hac, T. Heimburg, and H. Grubmüller. *Biophys. J. 85*, 1647-1655 (2003).
47. P. Mukhopadhyay, L. Monticelli, and D. P. Tieleman. *Biophys. J. 86*, 1601-1609 (2004).
48. G. Coadou, N. Evrard-Todeschi, J. Gharbi-Benarous, R. Benarous, and J.-P. Girault. *C. R. Acad. Sci. Paris, Chemie/Chemistry 4*, 751-758 (2001).
49. P. Henklein, R. Kinder, U. Schubert, and B. Bechinger. *FEBS Lett. 482*, 220-224 (2000).
50. S. Zheng, J. Strzalka, C. Ma, S. J. Opella, B. M. Ocko, and J. K. Blasie. *Biophys. J. 80*, 1837-1850 (2001).
51. W. B. Fischer, L. R. Forrest, G. R. Smith, and M. S. P. Sansom. *Biopolymers 53*, 529-538 (2000),
52. V. Wray, R. Kinder, T. Federau, P. Henklein, B. Bechinger, and U. Schubert. *Biochemistry 38*, 5272-5282 (1999).

Date Received: September 17, 2005

Communicated by the Editor Ramaswamy H. Sarma