Minireview

Vpu from HIV-1 on an atomic scale: experiments and computer simulations

W.B. Fischer*

Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK

Received 28 March 2003; revised 12 May 2003; accepted 13 May 2003

First published online 17 July 2003

Edited by Maurice Montal

Abstract Vpu is an 81 amino acid protein encoded by HIV-1. Its role is to amplify viral release by two mechanisms: (i) docking to CD4 with the consequence of targeting CD4 for ubiquitine-mediated degradation, and (ii) formation of ion channels to enhance particle release. The intensive research on its in vivo function, combined with structural investigations, makes this viral membrane protein one of the better characterised membrane proteins. The wealth of structural information enables the use of computational methods to elucidate the mechanisms of function on an atomic scale. The discovery of Vpu and the development of structural models in a chronological order is summarised and first efforts on investigating the mechanics are outlined.

Key words: HIV-1; Vpu; Viral membrane protein; Protein structure and function

1. Introduction

The viral genome of HIV-1 encodes a series of essential enzymes such as reverse transcriptase, protease and integrase [1]. Besides these principal proteins, a large number of non-structural proteins such as Tat (trans-activator), Rev (regulator of expression of virion protein), Nef (negative factor), Vif (virion infectivity), Vpr and Vpu (virus protein R and U, respectively) are encoded by the RNA of the virus [2]. The latter four proteins are called auxiliary proteins.

Vpu has been found to have a dual role, one which is achieved via protein–protein interaction, the other via a function as an ion channel. Other viruses also encode short membrane proteins of around 100 amino acids for which ion channel activity is established or at least suggested (e.g. M2 from influenza A [3,4], NB from influenza B [5], 6K protein from alphavirus [6], p7 from Hepatitis C [7]).

Vpu is a highly conserved protein. Strains from several isolates show a high level of sequence identity and homology over the entire sequence (Fig. 1), making the protein a potential drug target. Recently, a derivative of amiloride has been found to suppress the function of Vpu as an ion channel [8]. Besides its important medical relevance, Vpu may also serve as a test case for biophysical investigations on structure–function correlation in general. It has taken just 15 years from its discovery to the generation of reasonably good structural models which can even be used for in silico drug screening and to investigate the role of individual amino acids on an atomic level.

The basis of this review is to report the gain of knowledge in a chronological order. The order demonstrates the important role of each technique in every step of the discovery of Vpu, its function and its mechanisms.

2. Discovery and function

In 1988 two groups independently identify an unknown (U) virus protein (VP) in the genome of HIV-1 with a length of 81 amino acids and a molecular weight of 9 kDa [9,10,12]. This protein, named Vpu, is unique to HIV-1 and not found in HIV-2 and SIV [10]. According to the hydrophobicity of the encoded amino acids, a single highly hydrophobic stretch of 27 amino acids at the N-terminus is proposed and the protein consequently defined as 'membrane-associated' [9] (Fig. 2). Following the N-terminal hydrophobic amino acids, the remaining amino acids towards the C-terminal end include 15 charged residues. The presence of Vpu leads to a 10-fold amplification of virus replication. However, the absence of Vpu does not affect the cytopathicity and the kinetics of viral replication. Within the cytoplasmic part Vpu contains two phosphorylation sites at two serine residues (Ser-52 and Ser-56) [11,12].

Vpu is involved in the enhancement of particle release [13]. This function involves two distinct effects: Vpu initiates the degradation of CD4 molecules in the endoplasmic reticulum via its cytoplasmic domain [14–16]. The phosphorylated residues (Ser-52 and Ser-56) on the cytoplasmic site seem to be non-essential for Vpu to bind to CD4. However, these residues are crucial for directing the Vpu–CD4 complex to the ubiquitin-dependent proteasome degradation pathway [17–20]. The regions responsible for the docking of Vpu to CD4 are found to include residues 28–47 [19] and residues at the C-terminal end [21], especially residues 76–81 [19].

Another function of Vpu has been discovered which is to enhance the release of particles from the infected cells [16]. This role is allocated to the transmembrane (TM) domain. The topology of Vpu is identified as to be a type-I integral membrane protein which is capable of oligomerisation into

*Fax: (44)-1865-275234.
E-mail address: wolfgang.fischer@bioch.ox.ac.uk (W.B. Fischer).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00782-8

```
                10        20        30        40        50        60        70
              *.  :  *.* *.** *:**: **:** **** :* :**:**  ::::: **** ******:*:  . *  : **
Vpu_HV1B8  - QPIPIVTIVA LAVAIIIAIV VWSIVIIEYR KILRQRKIDR LIDRLIERAE DSGNESEGEI SALVEMGVEM
Vpu_HV1H2  - QPIPIVAIVA LVVAIIIAIV VWSIVIIEYR KILRQRKIDR LIDRLIERAE DSGNESEGEI SALVEMGVEM
Vpu_HV1B1  M QPIQIA-IVA LVVAIIIAIV VWSIVIIEYR KILRQRKIDR LIDRLIERAE DSGNESEGEI SALVEMGVEM
Vpu_HV1LW  M QPIQIA-IVA LVVAIIIAIV VWSIVIIEYR KILRQRKIDR LIDRLIERAE DSGNESEGEI SALVEMGVEM
Vpu_HV1BR  M QPIQTA-IAA LVVAIIIAIV VWSIVIIEYR KILRQRKIDR LIDRLIERAE DSGNESEGEI SALVEMGVEM
Vpu_HV1EL  M QPIGIIATAA LVVAIIIAIV VWTIVFIEYR RIKKQRRIDC LLDRITERAE DSGNESEGDR EKLSKL-VEM
Vpu_HV1Z2  M QPSQIIATAA LVVAAIIAIV VWTIVFIEYR RIKRQRRIDC IIDRIRERAE DSGNESEGDR EELSKL-VEM
Vpu_HV1S1  M QPLQIIAIVA LVVAAIIAIV VWTIVYIEYR KILRQRRIDR LIDRITERAE DSGNESEGDQ EELSAL-VER
Vpu_HV1Y2  L QSLQVLAIVA LVVATIIAIV VWTIVFIEYR KILRQRRIDR LINRITERAE DSGNESEGDQ EELSAL-VER
Vpu_HV1JR  M QPLQILAIVA LVVAGIIAII VWSIVLIEYR KILRQRRIDR LIDKIRERAE DSGNESEGDQ EELSAL-VER

                80
              ** ****:;* *
Vpu_HV1B8 GHHAPWDVDD L
Vpu_HV1H2 GHHAPWDVDD L
Vpu_HV1B1 GHHAPWDVDD L
Vpu_HV1LW GHHAPWDVDD L
Vpu_HV1BR GHHAPWDIDD L
Vpu_HV1EL GHHAPWDIDD L
Vpu_HV1Z2 GHHAPWDIDD L
Vpu_HV1S1 GHLAPWDVDD L
Vpu_HV1Y2 GHLAPWDVDD L
Vpu_HV1JR GHLAPWDIND L
```

Fig. 1. Multiple sequence alignment (Clustal X (1.82)) of 10 strains of Vpu. '*' denotes fully conserved residues, ':' denotes strong similarity, '.' stands for weak similarity and '-' indicates a gap. Since not all the strains start with methionine (M) or leucine (L) the first residue is set to glutamine (Q). Regions with series of fully conserved residues appear from residue Ile-8 to Asp-39, Glu-47 to Gly-58, and Val-68 to Leu-81. These regions do not necessarily define helical regions (Table 1) but also include interconnections between helical segments. Trp-22 and the EYR motif (Glu-28, Tyr-29, Arg-30) are completely conserved, whilst Ser-23 can be exchanged by a threonine. It may be assumed that with this high sequence identity all Vpu proteins adopt the same secondary and tertiary structure. Thus, the structure-based function of Vpu allows very little compromise.

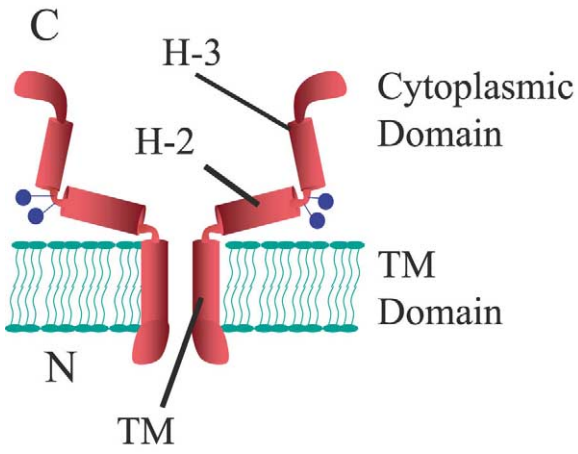

Fig. 2. Compromised sketch of the putative structure of a Vpu protein. Blue balls denote the phosphorylation sites at Ser-52 and Ser-56. Helix-1 at the N-terminal end (N) is often referred to as TM helix. H-2 and H-3 are helix-2 and helix-3, respectively. The remaining parts towards the C-terminal end (C) include either a short turn or a helix. For down-regulation of CD4 Vpu acts probably as a monomer, whilst for enhancing particle release it acts by forming an ion channel. The quaternary structure of the cytoplasmic form is unknown whether it is for the form in a monomer or as a homo-oligomer. Ion channels are formed by a self-assembly process of the monomers within the TM region. In this figure just two opposing Vpu proteins defining a pore are shown.

higher molecular-weight structures of up to 80 kDa in the presence of membranes and 85 kDa in the absence of membranes [22]. These values suggest at least a tetramer as a putative oligomerisation state in vivo. The results imply that one of the functions is to alter the intracellular environment probably via ion channel formation. Vpu expressed in *Escherichia coli* [23] or *Xenopus* oocytes [24] and reconstituted into lipid bilayers indeed induces the formation of an ion channel with preference for cations [23] and a minor indication of permeability of divalent cations [24]. Proton conductance via a fluorescence-quenching experiment is not detected [23]. Channel activity is found to be due solely to the TM region of Vpu; channel activity of a peptide corresponding to the TM segment and reconstituted into lipid bilayers has been found [16]. Other studies with Vpu expressed in *Xenopus* oocytes reveal a decrease in membrane conductance which is found to be associated with the capability of Vpu to down-modulate membrane protein expression [25].

It is also suggested that Vpu mediates the efficient membrane association of other proteins with the consequence of enhancing the particle release [26]. Expression of Vpu in BHK cells leads to an increase of permeability of the cells to hydrophilic molecules, leading to the suggestion that Vpu may also act as a viroporin [27]. However, the ultimate effect of a viroporin, leading to cell death by uncontrolled increase of permeability, has not been found so far [26].

Mutants within the TM region [28] as well as deletions of some of the hydrophobic amino acids within this region [29] seem to reduce the capability of Vpu to enhance the release of virus particles. Since the mutants show normal translocation patterns and are capable of oligomerisation with wild-type Vpu it is supposed [29] that structural integrity is another key factor for the TM segment to fulfil its role. A single mutant, in which Trp-22 is shifted two positions towards the C-terminal end (replacing Ile-24), is not able to induce CD4 degradation, indicating that there might not be a strict separation in mechanism of function between the two domains [28].

The mechanisms of Vpu function are via protein–protein interaction and ion channel activity. Both functions seem to be independently allocated with either one of two distinct structural elements: the cytoplasmic domain or the TM domain.

3. Spectroscopic investigations

The discovery that the two functional roles are attributed to the two structural domains provides the opportunity to per-

Table 1
Proposed secondary structural elements of Vpu and their putative lengths, including the techniques used to derive the data

	TM helix (helix-1)	Helix-2	Helix-3	Turn (T)/Helix (H)	Reference
NMR/CD		42–50	57–69	74–77 (T)	[30]
CD	5/6–29				[35]
NMR		37–51	57–72	73–78 (T)	[31]
NMR		40–50	60–68	75–79 (H)	[33]
SSNMR/NMR	9–29	30–49	58–70	–	[34]
MD simulations	5–25				[39]

Positions for the residues that define the start and the end of the secondary structural element have been taken from the references. NMR stands for solution NMR, SSNMR denotes solid state NMR.

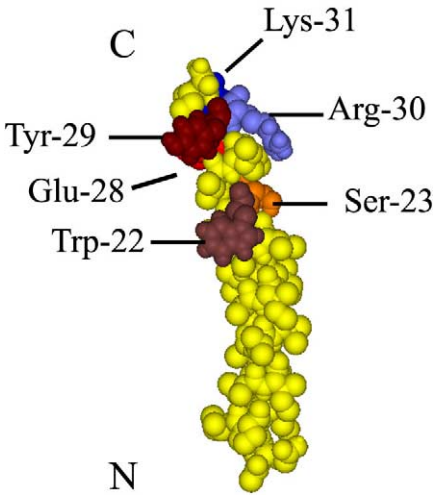

Fig. 3. Single TM helix (Vpu_{1-32}: MQPIPVAIV[10]ALVVAIII-AI[20]VVWSIVIIEY[30]RK) in Van der Waals representation (DSViewerPro50) after 3 ns of molecular dynamics simulations in a fully hydrated lipid bilayer. In this simulation 128 lipid molecules (1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phosphatidyl-choline, POPC) are surrounding the peptide. The system is then embedded in a water bath containing ca. 40 water molecules per lipid and several counter ions to obtain an uncharged box for the simulation. In total the simulation is undertaken with around 20 000 atoms. In the figure lipids and water molecules are omitted. Tyrosine and tryptophan residues are highlighted in dark and light brown, respectively, serine in orange and arginine in blue. Glu-28 (red) and Lys-31 (dark blue) are covered by the helix in this perspective. All other residues, mainly hydrophobic, are shown in yellow, including Gln-1, Pro-2 and Pro-4. The section between Trp-22 and Arg-30 defines the part of the protein which is embedded in the phospholipids headgroup region. The helix adopts a considerable kink and tilt. C and N denote the C- and N-terminal ends, respectively.

form structural analysis with each of the domains and even parts of them individually.

One of the earliest spectroscopic investigations attempting to achieve a structural model on the cytoplasmic site has been done by circular dichroism (CD) spectroscopy [30]. Nine overlapping 15 amino acid fragments starting from Vpu_{28-42} to Vpu_{67-81} were synthesised using standard solid-phase peptide synthesis (SPPS). The frames overlapped to avoid any erroneous interpretation of the data due to chain end effects. For three constructs, Vpu_{41-58}, Vpu_{52-74}, Vpu_{63-81}, ^1H nuclear magnetic resonance (NMR) spectra were recorded, revealing a helical motif for residues 42–50 (Vpu_{41-58}) and 57–69 (Vpu_{52-74}) (Table 1). The C-terminal end includes a reversed turn for residues 74–77 (Vpu_{63-81}).

Studies have then been extended using longer constructs: Vpu_{32-81} also synthesised by SPPS [31]. Solution NMR data recorded in distilled water (H_2O or D_2O) containing 50% TFE-d_2 by volume and processed with Xplor 3.1 can be summarised as follows: a helical motif from residues 37–51, an interconnecting loop from residue 52–56, a second helix from residue 57–72 followed by a turn motif from residues 73–78 (Table 1). In a non-bilayer environment, replacement of the two phosphorylation sites (Ser-52 and Ser-56) by asparagines seems not to affect the overall structure in these experiments. However, for a Vpu_{41-62} construct it has been found by solution NMR that the absence of the phosphorus groups destabilises the helical motif [32].

Solution NMR spectra using the cytoplasmic domain of Vpu dissolved in an electrolyte solution (100 mM NaCl) reveals a similar secondary structural arrangement [33] as described above [31]. The data suggest a well-defined amphiphilic helix between residues 40 and 50 and a less well-defined helix between residues 60 and 68. Towards the C-terminal end, another short helix is formed by residues 75–79 (Table 1). The Vpu construct used is an expression product from BL21(DE33) cells carrying the pET16b plasmid.

Solution NMR on two Vpu constructs, Vpu_{2-37} and Vpu_{28-81} from *E. coli* reconstituted into vesicles, defines the sequence of the secondary structural elements as the following: a 20 amino acid helix from residues 30–49, a linker region from residues 51–57, a helix from residues 58–70 (13 residues), followed by a C-terminal region, which is not further specified [34].

Despite the apparent similarity of the investigations regarding the secondary structural elements, the constructs vary in tertiary structure. Whilst studies in H_2O/TFE [31] and in vesicles [34] reveal a more extended cytoplasmic domain, the study undertaken in an electrolyte solution suggests a more compact arrangement of the helices [33].

The TM segment of Vpu reveals a helical motif characterised mostly by solid state NMR (SSNMR) spectroscopy [35,36] (Fig. 3). Spectra obtained by SSNMR of uniformly ^{15}N-labelled full-length Vpu, a truncated construct including the TM domain and the amphiphatic helix-2 (Vpu_{2-51}), as well as a construct of purely the cytoplasmic domain (Vpu_{28-81}), show that the first two constructs reveal a helix perpendicular to the membrane surface with a tilt angle of approximately 15° [36]. These measurements establish that the second and the third helix are aligned parallel to the membrane surface. All constructs are derived from expression systems.

SSNMR spectra of solely TM constructs of Vpu (Vpu_{1-39}) derived from SPPS with labelled alanine residues are in agreement with an overall tilt of a helix close to 20° [35]. It is suggested that the helical motif also includes the aromatic residues Trp-22 and Tyr-29. Solution NMR data on Vpu^{1-39} reveal a stable folded tertiary structure which can be summarised as a U-shaped fold with an N-terminal turn (1–6), a linker (7–9), short helix (10–16) and a loop to another helix extending from residue 22–36. A bend is found at Tyr-29, which leads to the conclusion that Trp-22 and Tyr-29 are within the TM helix. It is suggested that the helical motif is not apparent before residue 5 or 6.

Solution NMR spectra monitoring H/D exchange on full-length Vpu expressed (*E. coli*) and reconstituted into micelles have revealed the first experimental evidence for the putative length of the TM helix [34]. Residues Ala-9 to Arg-29 show hindered exchange.

Recent investigations have used SSNMR [37] on lipid bilayer systems or synchroton radiation-based X-ray reflectivity measurements on lipid monolayers [38] to analyse the association of helix-3 with the bilayer. It is suggested that helix-3 is loosely bound to the lipid bilayer [37] or at least its contact with the surface of a bilayer is dependent on low lateral pressure of the bilayer [38]. Whilst the first study is achieved with a synthetic Vpu construct (Vpu_{51-81}), the latter is performed with full-length Vpu. In both investigations a backpack formation of helix-3 riding on top of helix-2 is suggested as a possible site of helix-3.

4. MD simulations bridge the gap towards investigations on an atomic scale

MD simulations have been performed on a bundle of parallel-aligned helices corresponding to the TM part of Vpu. The length of the helix is defined by the number of hydrophobic residues present in the sequence. Recent molecular dynamic simulations on single helices embedded in a fully hydrated lipid bilayer suggest the length of the TM helix to be from residues 5 to 25 [39].

Channel activity of Vpu has been investigated by generating a putative pore with assembled parallel TM helices (bundles) based on the sequence Ac-IVAIVALVVAIIIAIVVWSIVII-NH_2 [40]. The helical motif is in accordance with findings for the TM domain of the viral proton channel M2 from influenza A based on NMR spectroscopic investigations at that time [41]. It is assumed that Vpu shares this topology because of the homology in amino acid length and its function (enhancement of viral release). Bundles of four, five and six TM helices are hydrated with explicit water molecules (TIP3P). The relative orientation of the segments to each other is such that all the serines are facing the putative pore of the bundle. The representation of the bilayer potential is based on a residue by residue hydrophobicity in the simulation protocol. All results emerge from 100 ps MD simulations using the Charmm force field. Restraints have been imposed on the water molecules at the two ends of the bundles to prevent evaporation and the TM helices have been held together so that they remain in a helical conformation. Long-range electrostatics have been truncated (cut off with shift function) beyond a radius of 1.3 nm. An assessment of ion conductivity based solely on the pore radius using the programme HOLE [42] has revealed about 60 pS for the pentameric bundle. Comparison with experimental results makes this assembly the most favourable model for a putative ion channel to date. Potential energy profiles of Na^+ and Cl^- within the pentameric pore support the experimental finding of a slight preference of the Vpu channel for cations, which is taken to further strengthen the argument for this bundle to be the most likely pore.

In another study simulations are performed with an unrestrained pentameric bundle embedded in a solvated (with explicit water molecules based on the flexible SPC/E model) bilayer mimetic slab of octane [43]. The model is based on the sequence: Ac-IVAIVALVVAIIIAIVVWSIVII-NH_2. Simulations have been run with a Charmm force field for up to 1 ns, using the Ewald method to account for long-range electrostatics in the simulation box. Models have been investigated with either Ser-24 or Trp-23 facing the pore. In both cases the bundles adopt a conical shape with a larger opening on the C-terminal end. All simulations uncover rotational motions of individual helices. The bundle based on the oriented Trp-23s shows kinks in two of the helices at the end of the simulation. In addition, no continuous water column was observed in either of the models. It is interesting to note that a tilt angle of 4.2° is reported with no supercoiling, in agreement with earlier simulations [40].

In a combined experimental and computational approach the number of helices forming a pore in vivo has been explored on an extended TM sequence: MQPIPVAIV^{10}ALV-VAIIIAI^{20}VVWSIVIIEY^{30}RK [44] (Fig. 4). For the simulations the Gromacs force field is used and the long-range electrostatics have been truncated beyond 1.4 nm (cut off). Bundles of four, five and six helices are embedded in a fully hydrated (SPC water model) lipid bilayer of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) molecules.

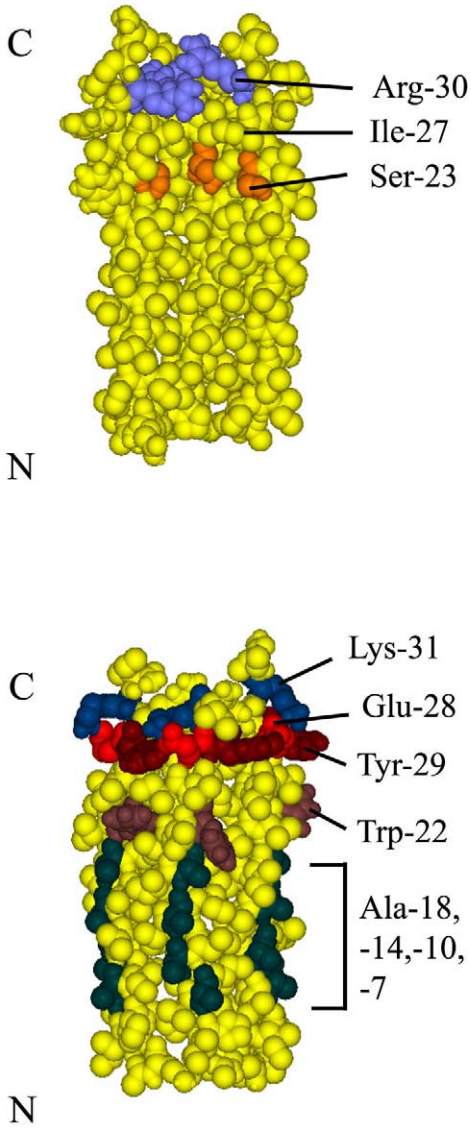

Fig. 4. A view inside the pore of a pentameric bundle of TM helices (Vpu_{1-32}) (top panel). The conditions for the simulations are identical to those mentioned in Fig. 3, except for a lower number of lipids (96 molecules) forming the bilayer. The lower number of lipids is necessary to allow the bigger peptide structure to be inserted into the lipid bilayer without increasing the size of the simulation box and the overall number of atoms for the simulation. For the display, two helices are omitted to obtain a view inside the pore. Arginines (Arg-30, blue) and serines (Ser-23, orange) form rings of hydrophilic residues, separated by a ring of isoleucines (Ile-27, yellow and indicated by the line). The pore towards the N-terminal and (N) is lined by purely hydrophobic residues. A view of the outside wall formed by the three helices is shown in the lower part of the figure (bottom panel). Tyrosines (dark brown) from one helix and the glutamates (red) and lysines (dark blue) from an adjacent helix are shown. They are forming a clamp-like arrangement, possibly stabilising the bundle assembly. In addition, the tryptophans (light brown) and the line of alanines (dark green) are highlighted. Their role in the function of Vpu can only be speculated as to be important for either protein–lipid contact or protein–protein interaction. C and N denote the C- and N-terminal end, respectively.

Conductance values calculated for the models, assuming that they are hydrated by a 0.5 M KCl solution, after a 1 ns simulation (tetramer: (12.2 ± 4.1) ps, pentamer: (52.2 ± 5.0) ps, hexamer: (51.9 ± 2.5) ps; Table 2), are in support of a pentameric or possibly a hexameric assembly, since experiments confirm the conductance of approximately 60 pS (recorded for a 0.5 M KCl solution [44]). The similar conductance found in the calculations for the pentamer and hexamer are caused by the partial occlusion of the pore by the side chain of Arg-30, which also points into the pore (Fig. 4, top panel). In the bundle, these arginines form a positive ring followed by a ring of isoleucines (Ile-27) and serines (Ser-23). In comparison, calculations for NB, an ion channel-forming protein of influenza B sharing approximately the same sequence length and topology as Vpu (see [45]), shows conductance of 22, 56, and 88 pS for the tetrameric (NB-4), pentameric (NB-5) and hexameric bundle (NB-6), respectively [46] (Table 2). Compared with experimental findings [5,46] the values for the tetramer do not exclude this bundle from its possible contribution to channel activity of NB. Calculations of the conductivity of the bundles calculated in [40] support the overall trend. However, they do not allow a direct comparison in this context since the water molecules have been restrained at the mouth of the pore.

In a tetrameric bundle of Vpu, hardly any water molecules (∼three; Table 2) are discovered at the end of the simulation, which is taken as an indication for excluding this assembly as a putative pore [44]. The pore narrows down to a minimum radius of about 0.7 Å for a stretch of about 20 Å towards the N-terminal end. In contrast, NB-4 still allows 16 times more water molecules (approximately 49; Table 2) within its pore. The pore radius in NB-4 gradually reduces towards the C-terminal end with a minimum value of 1.1 Å for a short stretch of approximately 5 Å. This radius would allow at least for single-filed water within this short narrow pass (see e.g. gramicidin A with a pore radius between 1.5 and 2.0 Å across the full length of the bilayer [47–50]). Thus, pore radius for Vpu-4 is definitively too small for water to be almost permanently present in this part of the pore. The length of this narrow pass of almost 20 Å makes it energetically more demanding to open under physiological conditions. Averaged tilt and kink angles of the Vpu bundles are calculated as 15.3°/14.5°/6.0° (tetramer/pentamer/hexamer) and 16.9°/19.9°/12.7° (tetramer/pentamer/hexamer), respectively (Table 2). The tilt angles seem to decrease with an increasing number of helices forming the bundle. For a hexamer, tilt angles of Vpu and NB are indistinguishable. Thus, for pentameric bundles the amino acid sequence strongly drives the packing of the bundle. It is important to note that one of the helices of the tetramer moves away from the bundle whilst another helix adopts an unusual kink of about 32.1°. Arg-30 at the C-terminal end of the bundle is responsible for the conical-like shape of the pore. Another reason for a putative pentameric structure of Vpu, responsible for ion conductance, might be the alignment of the glutamates and lysines of one helix with the tyrosines on the adjacent helix (Fig. 4, bottom panel). These residues form a clamp based on favourable electrostatic and hydrogen-bonding interactions. The role of the line of alanines (Fig. 4, bottom panel) needs still to be elucidated.

Simulations on a single helix in a lipid bilayer reveal a kink of about 15° [39]. Thus, the TM segment of Vpu might not be a straight helix, which is indicated also experimentally by solution NMR [35]. A reason for this might be the long hydrophobic stretch towards the N-terminal end of the bundle which allows for a more tidy association of the helices compared to the C-terminal part with its serines and arginines. This hydrophobic stretch might form a hydrophobic lock or gate for channel activity.

Simulations using both an octane slab to mimic the lipid bilayer and an explicit lipid bilayer also suggest that the pentameric assembly is the most likely form of a Vpu channel [51]. Large translational movement of an individual helix within the hydrophobic octane slab is taken as an indication of the repulsion of a sixth helix in favour of the pentameric assembly. In addition, water seems to fill only the N-terminal half of the bundle. The bundle has been built such that the hydrophilic residue serine is facing the hydrophobic part of the bilayer rather than the lumen of the pore. This leaves the tryptophans at the helix/helix/bilayer interface in the initial starting structure.

5. What is left to be done?

5.1. Is there a gating?

The role of tryptophans in the mechanism of Vpu function has been investigated using MD simulations in a fully hy-

Table 2
Structural and functional features of tetrameric (Vpu-4, NB-4), pentameric (Vpu-5, NB-5) and hexameric bundles (Vpu-6, NB-6) of Vpu and the viral ion channel forming protein NB from influenza B [46]

	Vpu-4	Vpu-5	Vpu-6	NB-4	NB-5	NB-6
Calculated conductivity (pS), 0.5 M KCl	12.2 ± 4.1	52.2 ± 5.0	51.9 ± 2.5	ca. 22	ca. 56	ca. 88
0.5 M NaCl*	21*	57*	93*			
No. of water molecules	∼3	∼92	∼108	∼49	∼105	∼145
	139*	257*	381*			
Minimum pore radius (Å)	0.7 ± 0.1	2.5 ± 0.5	2.3 ± 0.6	ca. 1.1	ca. 2.8	ca. 3.9
	ca. 0.7*	ca. 1.5*	ca. 2.5*			
Kink angle (°)	16.9 ± 10.0	19.9 ± 7.8	12.7 ± 7.7	20.6 ± 6.8	13.8 ± 5.9	8.8 ± 4.04
	(32.1 ± 1.9)					
Tilt angle (°)	15.3 ± 5.0	14.5 ± 3.4	6.0 ± 2.1	4.8 ± 1.3	9.2 ± 2.8	6.2 ± 1.8

Values with standard deviation derive from averaging the last 250 ps in steps of 50 ps of the trajectory. Conductivity and minimum pore radius are calculated using the programme HOLE [42], which does not take electrostatics into account and may be used for an assessment of pore conductivity. Conductivity is scaled by a factor of 0.5 [44] and 0.2 for values denoted by '*' [40]. Kink and tilt angles are also averages over the number of helices for each bundle. The number of water molecules in each bundle is estimated from the structure after 1 ns of simulation. For Vpu-4 a single helix out of the four shows a large kink angle, which is shown in brackets. Values indicated by '*' are obtained by Grice et al. [40]. Values for which no standard deviations are available are indicated by 'ca.'.

drated lipid bilayer using the Gromacs package [52]. Simulations have been run with a pentameric bundle based on experimental findings in combination with a global search protocol [53], which suggest that the energetically favourable bundle is the one in which the TM helices of Vpu (AIVALVVAIIIAIVVWSIVIIE) have all tryptophans facing the pore. The data are compared to those with newly generated bundles using a simulated annealing protocol combined with short MD simulations (SA/MD) using the programme XPLOR. Whilst the bundle based on experimental data is falling apart, the computationally derived bundle retains a bundle-like assembly. Thus, having tryptophans facing the pore might be a putative closed form of the pore and in accordance with other findings of rotational-like motion, a gating mechanism could be due to rotation of one or more helices, thereby moving the Trp-22 into the pore to achieve closure.

5.2. *What is the role of other parts of Vpu and do they affect the TM helix?*

To address these questions a larger model of Vpu is generated (using XPLOR) including the first 52 amino acids (Vpu_{1-52}) [54] (Fig. 5). The model is built by generating a helix including the 52 residues and bending it around residues Glu-28 to Ile-32, which are assumed to form the bend so that the C-terminal end comes to align as a helix with its long axis parallel to the membrane surface (Fig. 5AI). Subsequent 9 ns simulations reveal a fairly stable helical region up to residue Tyr-29 (Fig. 5BI). According to the ϕ and ψ angles Arg-30 and Lys-31 and Ile-32 strongly deviate from a helical conformation. Helix-2 is rolling over, enabling Glu-28, Lys-31 and Arg-34 to form a complex salt bridge. Helix-2 remains in a helical environment because of its unique position in the bilayer. A ridge of alternating positively and negatively charged amino acids (Asp (red)/Arg (blue) and Glu (red)/Arg (blue), Fig. 5) in helix-2 form another extended complex salt bridge, whereas the underlying hydrophobic part (isoleucines, leucines) is embedded in the hydrophobic slab of the lipid bilayer (Fig. 5AII,BII). The tilt angle of the TM helix versus the membrane normal (23.3° ± 2.1°) is in support of the experimental findings from NMR spectroscopy. Also, a considerable kink (from 7° to sometimes 15°) in the TM helix is observed in the simulations. The TM helix and helix-2 span an approximate angle of 100°. As a conclusion, only parts of the EYR motif (Glu-28, Tyr-29, Arg-30) are involved in the bend. Helix-2 seems to be a fairly stable unit that can be described as a peptide float functioning as a spacer, separating the TM helix from the C-terminal part. Thus, Vpu is a construct of a series of functional modules: an ion channel-active TM helix linked via a clamp (Glu-28, Lys-31, Arg-34) to a rigid peptide float (helix-2) acting as spacer between the TM helix and the rest of the cytoplasmic domain beyond helix-2.

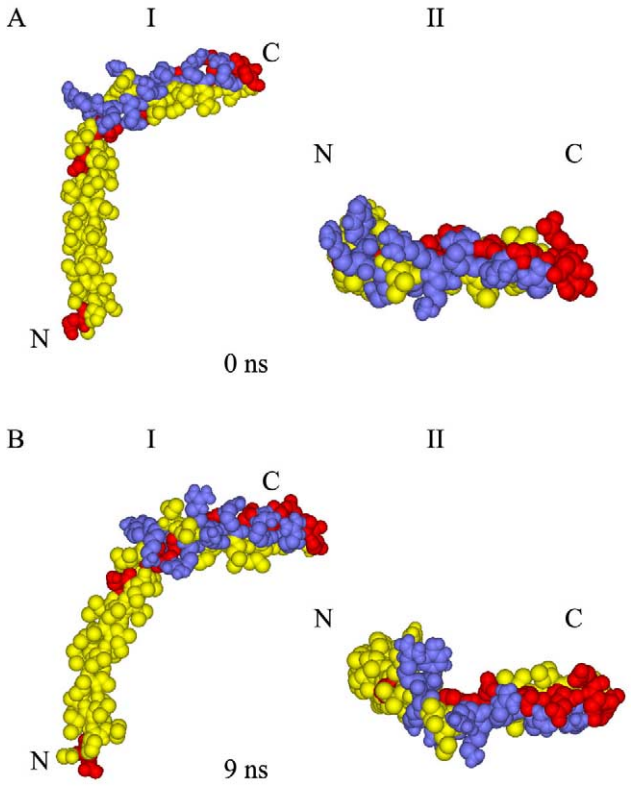

Fig. 5. Side view (left side (I)) and view on top of helix-2 (right side (II)) of the kinked model of Vpu (Vpu_{1-52}) at the beginning (panel A) and end (panel B) of a 9 ns MD simulation. All hydrophilic residues are highlighted in blue (basic residues) and red (acidic residues) including Gln-1, Ser-23 (all red) within the TM helix and Ser-52 in helix-2. All the hydrophobic residues are shown in yellow. The bend has been accomplished around residues Glu-28 to Ile-32 (interconnecting region) so that the residues are held as best as possible in a helical conformation [54]. This places helix-2 so that it comes to lie on top of the bilayer. Ser-23 of the TM helix faces away from helix-2. The rationale behind this positioning is that in an oligomeric assembly of parallel TM helices these serines are facing into a pore rather than pointing under helix-2. When placed into a lipid bilayer more lipid molecules have been taken away from the leaflet which embeds helix-2 than for the leaflet which harbours the N-terminal end of the TM helix. To close the gap underneath helix-2, 600 ps of equilibration have been undertaken with the peptide position restrained, which allows the hydrophobic tails of the lipid molecules to crawl underneath helix-2 before the whole system is used for the simulation (production phase) (panels AI and AII). After the 9 ns simulation, the TM helix is considerably kinked (panel BI) and helix-2 has undertaken an almost 90° revolution accompanied by an unwinding of the interconnecting region (panel BII). The hydrophilic residues on helix-2 define a ridge based on an alternating aspartate and glutamate to either arginine or lysine repeat (*i*,*i*+1 and *i*,*i*+3), which forms a complex salt bridge stabilising the helix structure (panels AII and BII). This arrangement leads to a a slight left-handed twist. C and N denote the C- and N-terminal end, respectively.

6. Other viral ion channels

M2, the viral proton channel from influenza A, is also a well-characterised viral membrane protein. Driven by the immediate medical relevance of this protein, detailed structural models of the TM domain have been generated using SSNMR spectroscopy [41,55] and refined by molecular dynamics simulations [56]. Plausible models for the mechanisms of function of this channel have been developed by computational simulations [57–59]. With its known channel blocker, amantadine, M2 comprises one of the best-characterised model to date. First attempts have been made to derive full-length M2 by solid-phase peptide synthesis [60], which makes it only a matter of time until more structural detail of its extramembraneous domains will emerge. In the meantime, for example sequence similarities and structural analogies with Vpu can be used to predict the structural motifs of the cytoplasmic domain of M2 [61].

All other viral ion channels have yet to reach an equivalent state where plausible experimental based structural models can be generated.

7. Conclusions

Our knowledge about Vpu seems to be more advanced compared to the structural information available on the M2 proton channel from influenza A. While there is a lot of information about the function of Vpu, so far structural information has enabled us to draw a moderately detailed structure and has provided enough information for performing MD simulations on an 'all atom' basis. As a final stage, and probably just as exciting and time-consuming, is the determination of the mechanism of Vpu activity and the roles played in it by the individual amino acids. In this respect we are just at the beginning of the journey. It may appear to be an unnecessary 'hunt for details'; nevertheless, these studies will give us a clue of the workload when dealing with more complex proteins or on a larger scale in biochemical proteomics to fully understand a whole array of proteins on an atomic scale.

Acknowledgements: The author acknowledges the E.P. Abraham Research Fund (Oxford, UK) for financial support. Thanks to A. Watts, J. Watts, S. Grage, V. Lemaitre, C.G. Kim, Y.H. Lam, and D. Fischer (Oxford) for valuable discussions.

References

[1] Greene, W.C. and Peterlin, B.M. (2002) Nat. Med. 8, 673–680.
[2] Cimarelli, A. and Darlix, J.-L. (2002) Cell. Mol. Life Sci. 59, 1166–1184.
[3] Wang, C., Takeuchi, K., Pinto, L.H. and Lamb, R.A. (1993) J. Virol. 67, 5585–5594.
[4] Lin, T. and Schroeder, C. (2001) J. Virol. 75, 3647–3656.
[5] Sunstrom, N.A., Prekumar, L.S., Prekumar, A., Ewart, G., Cox, G.B. and Gage, P.W. (1996) J. Membr. Biol. 150, 127–132.
[6] Melton, J.V., Ewart, G.D., Weir, R.C., Board, P.G., Lee, E. and Gage, P.W. (2002) J. Biol. Chem. 277, 46923–46931.
[7] Griffin, S.D.C., Beales, L.P., Clarke, D.S., Worsfold, O., Evans, S.D., Jäger, J., Harris, M.P.G. and Rowlands, D.J. (2003) FEBS Lett. 535, 34–38.
[8] Ewart, G.D., Mills, K., Cox, G.B. and Gage, P.W. (2001) Eur. Biophys. J. 31, 26–35.
[9] Strebel, K., Klimkait, T. and Martin, M.A. (1988) Science 241, 1221–1223.
[10] Cohen, E.A., Terwilliger, E.F., Sodroski, J.G. and Haseltine, W.A. (1988) Nature 334, 532–534.
[11] Strebel, K., Klimkait, T., Maldarelli, F. and Martin, M.A. (1989) J. Virol. 63, 3784–3791.
[12] Schubert, U., Schneider, T., Henklein, P., Hoffmann, K., Berthold, E., Hauser, H., Pauli, G. and Porstmann, T. (1992) Eur. J. Biochem. 204, 875–883.
[13] Klimkait, T., Strebel, K., Hoggan, M.D., Martin, M.A. and Orenstein, J.M. (1990) J. Virol. 64, 621–629.
[14] Willey, R.L., Maldarelli, F., Martin, M.A. and Strebel, K. (1992) J. Virol. 66, 7193–7200.
[15] Kimura, T., Nishikawa, M. and Ohyama, A. (1994) J. Biochem. 115, 1010–1020.
[16] Schubert, U., Bour, S., Ferrer-Montiel, A.V., Montal, M., Maldarelli, F. and Strebel, K. (1996) J. Virol. 70, 809–819.
[17] Bour, S., Schubert, U. and Strebel, K. (1995) J. Virol. 69, 1510–1520.
[18] Friborg, J., Ladha, A., Göttlinger, H., Haseltine, W.A. and Cohen, E.A. (1995) J. Acquired Immune Defic. Syndr. Hum. Retrovirol. 8, 10–22.
[19] Margottin, F., Benichou, S., Durand, H., Richard, V., Liu, L.X. and Benarous, R. (1996) Virology 223, 381–386.
[20] Tiganos, E., Yao, X.-J., Friborg, J., Daniel, N. and Cohen, E.A. (1997) J. Virol. 71, 4452–4460.
[21] Chen, M.-Y., Maldarelli, F., Martin, M.A. and Strebel, K. (1993) J. Virol. 67, 3877–3884.
[22] Maldarelli, F., Chen, M.Y., Willey, R.L. and Strebel, K. (1993) J. Virol. 67, 5056–5061.
[23] Ewart, G.D., Sutherland, T., Gage, P.W. and Cox, G.B. (1996) J. Virol. 70, 7108–7115.
[24] Schubert, U., Ferrer-Montiel, A.V., Oblatt-Montal, M., Henklein, P., Strebel, K. and Montal, M. (1996) FEBS Lett. 398, 12–18.
[25] Coady, M.J., Daniel, N.G., Tiganos, E., Allain, B., Friborg, J., Lapointe, J.-Y. and Cohen, E.A. (1998) Virology 244, 39–49.
[26] Deora, A., Spearman, P. and Ratner, L. (2000) Virology 269, 305–312.
[27] Gonzalez, M.E. and Carrasco, L. (2001) Virology 279, 201–209.
[28] Tiganos, E., Friborg, J., Allain, B., Daniel, N.G., Yao, X.-J. and Cohen, E.A. (1998) Virology 251, 96–107.
[29] Paul, M., Mazumder, S., Raja, N. and Jabbar, M.A. (1998) J. Virol. 72, 1270–1279.
[30] Wray, V., Federau, T., Henklein, P., Klabunde, S., Kunert, O., Schomburg, D. and Schubert, U. (1995) Int. J. Peptide Protein Res. 45, 35–43.
[31] Federau, T., Schubert, U., Floßdorf, J., Henklein, P., Schomburg, D. and Wray, V. (1996) Int. J. Peptide Protein Res. 47, 297–310.
[32] Coadou, G., Evrard-Todeschi, N., Gharbi-Benarous, J., Benarous, R. and Girault, J.-P. (2002) Int. J. Biol. Macromol. 30, 23–40.
[33] Willbold, D., Hoffmann, S. and Rösch, P. (1997) Eur. J. Biochem. 245, 581–588.
[34] Ma, C. et al. (2002) Protein Sci. 11, 546–557.
[35] Wray, V., Kinder, R., Federau, T., Henklein, P., Bechinger, B. and Schubert, U. (1999) Biochemistry 38, 5272–5282.
[36] Marassi, F.M., Ma, C., Gratkowski, H., Straus, S.K., Strebel, K., Oblatt-Montal, M., Montal, M. and Opella, S.J. (1999) Proc. Natl. Acad. Sci. USA 96, 14336–14341.
[37] Henklein, P., Kinder, R., Schubert, U. and Bechinger, B. (2000) FEBS Lett. 482, 220–224.
[38] Zheng, S., Strzalka, J., Ma, C., Opella, S.J., Ocko, B.M. and Blasie, J.K. (2001) Biophys. J. 80, 1837–1850.
[39] Fischer, W.B., Forrest, L.R., Smith, G.R. and Sansom, M.S.P. (2000) Biopolymers 53, 529–538.
[40] Grice, A.L., Kerr, I.D. and Sansom, M.S.P. (1997) FEBS Lett. 405, 299–304.
[41] Kovacs, F.A. and Cross, T.A. (1997) Biophys. J. 73, 2511–2517.
[42] Smart, O.S., Neduvelil, J.G., Wang, X., Wallace, B.A. and Sansom, M.S.P. (1996) J. Mol. Graph. 14, 354–360.
[43] Moore, P.B., Zhong, Q., Husslein, T. and Klein, M.L. (1998) FEBS Lett. 431, 143–148.
[44] Cordes, F.S., Tustian, A., Sansom, M.S.P., Watts, A. and Fischer, W.B. (2002) Biochemistry 41, 7359–7365.
[45] Fischer, W.B. and Sansom, M.S.P. (2002) Biochim. Biophys. Acta 1561, 27–45.
[46] Fischer, W.B., Pitkeathly, M., Wallace, B.A., Forrest, L.R., Smith, G.R. and Sansom, M.S.P. (2000) Biochemistry 39, 12708–12716.
[47] Roux, B. and Karplus, M. (1991) Biophys. J. 59, 961–981.
[48] Woolley, G.A. and Wallace, B.A. (1992) J. Membr. Biol. 129, 109–136.
[49] Chiu, S.-W., Subramaniam, S. and Jakobsson, E. (1999) Biophys. J. 76, 1929–1938.
[50] de Groot, B.L., Tieleman, D.P., Pohl, P. and Grubmüller, H. (2002) Biophys. J. 82, 2934–2942.
[51] Lopez, C.F., Montal, M., Blasie, J.K., Klein, M.L. and Moore, P.B. (2002) Biophys. J. 83, 1259–1267.
[52] Cordes, F., Kukol, A., Forrest, L.R., Arkin, I.T., Sansom, M.S.P. and Fischer, W.B. (2001) Biochim. Biophys. Acta 1512, 291–298.
[53] Kukol, A. and Arkin, I.T. (1999) Biophys. J. 77, 1594–1601.
[54] Sramala, I., Lemaitre, V., Faraldo-Gomez, J.D., Vincent, S., Watts, A. and Fischer, W.B. (2003) Biophys. J. 84, 3276–3284.
[55] Wang, J., Kim, S., Kovacs, F. and Cross, T.A. (2001) Protein Sci. 10, 2241–2250.

[56] Forrest, L.R., DeGrado, W.F., Dieckmann, G.R. and Sansom, M.S.P. (1998) Fold. Des. 3, 443–448.
[57] Schweighofer, K.J. and Pohorille, A. (2000) Biophys. J. 78, 150–163.
[58] Forrest, L.R., Kukol, A., Arkin, I.T., Tieleman, D.P. and Sansom, M.S.P. (2000) Biophys. J. 78, 55–69.
[59] Smondyrev, A.M. and Voth, G.A. (2002) Biophys. J. 83, 1987–1996.
[60] Kochendörfer, G.G., Salom, D., Lear, J.D., Wilk-Orescan, R., Kent, S.B.H. and DeGrado, W.F. (1999) Biochemistry 38, 11905–11913.
[61] Saldanha, J.W., Czabotar, P.E., Hay, A.J. and Taylor, W.R. (2002) Prot. Pept. Lett. 9, 495–502.