Protein−Protein Interactions: Modeling the Hepatitis C Virus Ion Channel p7

George Patargias,[†] Nicole Zitzmann,[‡] Raymond Dwek,[‡] and Wolfgang B. Fischer*[,†,§]

Biomembrane Structure Unit and Oxford Glycobiology Institute, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, U.K., and Bionanotechnology Interdisciplinary Research Collaboration, Clarendon Laboratory, Department of Physics, Oxford University, Parks Road, Oxford OX1 3PU, U.K.

Received July 27, 2005

The p7 protein is a small ion-channel-forming membrane polypeptide encoded by the hepatitis C virus which consists of two transmembrane α-helices, TM1 and TM2, and can be blocked by long-alkyl-chain iminosugar derivatives. The length of TM1 and TM2 was estimated by employing different secondary structure prediction algorithms and is proposed to span from Ala-10 to Leu-32 for TM1 and from Trp-36 to Pro-58 for TM2. A configurational search protocol based on simulated annealing combined with short restrained molecular dynamics simulations is used in addition to protein−protein docking to investigate the packing of TM1/TM2. Full p7 oligomeric bundles were generated, and in the most plausible models serines and threonines are facing the hydrophilic pore. In these models, His-17 would be a pore-facing residue, suggesting that p7 may be sensitive to pH in respect to its function.

Introduction

Hepatitis C virus is the major cause of chronic viral hepatitis which can lead eventually to cirrhosis and liver cancer. The single-stranded RNA genome encodes a large polyprotein precursor of ∼3000 amino acids that is cleaved into 10 mature structural and nonstructural proteins (E1, E2, p7, NS2−NS5) by viral or host proteases. p7 is a small, transmembrane protein located between the structural and the nonstructural regions of the polyprotein. It has been shown to be essential for virus infectivity[1] but without affecting RNA replication. When reconstituted in artificial lipid bilayers, p7 forms ion channels[2] that are cation-selective at normal pH with measured conductances of between ∼86 and ∼100 pS.[3] Drugs such as amantadine,[4] hexamethylene amiloride,[2] and long-alkyl-chain iminosugar derivatives[3] block the activity of the channel, which makes p7 a potential antiviral target.

Previous analysis of p7 sequences extracted from different HCV isolates have suggested that the p7 polypeptide is composed of two long hydrophobic transmembrane (TM) stretches connected by a short basic cytosolic loop.[5] Mutations in the loop region abrogate the channel activity of p7.[6] The hydrophobic TM stretches are proposed to be α-helical based on secondary structure prediction results[5] and the functional similarity of p7 to other viral channels such as M2, NB, and Vpu, for which the helicity of the TM stretch has been experimentally verified.[7−12] However, there is almost no sequence homology between the individual strains (Figure 1A). There is no structural information available for p7, either in relation to its oligomerization state or the packing of the constituent helices. On the basis of electron microscopy and biochemical data, the channel is proposed to be made by a hexameric bundle with a putative pore diameter of 3−5 nm.[4] In the same paper, computer modeling was performed in order to evaluate the stability of the hexamer. By mapping the amino acids of p7 onto a backbone of two α-helical domains of bacteriorhodopsin, a structure of a channel with a 2.3 nm pore diameter was predicted. In this model, the amphipathic channel-forming N-terminal helix (TM1) is stabilized by the C-terminal helix (TM2) and the adjacent N-terminal helix of the next monomer.

Different computational techniques including global search methods,[13−17] Monte Carlo simulations, either with an all-atom force field[18] or with a simplified interaction[19,20] potential function, and protein−protein docking[21] have been developed to model and evaluate helix−helix interactions in pairs of helices. When combined with mutagenesis data[13] or with silent mutation analysis,[17] global search methods produced the successful prediction of helical dimers for glycophorin A and the CD-3 ζ, respectively. Helix−helix docking is a rational way of mimicking the hierarchical nature of the folding of membrane proteins since it models a process analogous to that of protein folding.[22] Helix−helix docking gave the correct prediction of the packing of the helix pairs 1−2, 2−3, and 4−5 of the L subunit of the photosynthetic reaction center.[21]

In the present study, a modeling procedure based on a combination of global search method, a simplified version of the method employed by Adams and co-workers,[13] and protein−protein docking was applied to generate most-probable models for p7 in the monomeric and the assembled form.

Results

Definition of the TM Helices. In total, 11 secondary structure prediction programs were used to predict the length of each helical stretch of p7 (Figure 1B). The length of the first helix (TM1) varies considerably across the results from the individual prediction methods. DAS and TMFINDER predict the helix to be as long as 13 amino acids. The other programs, ALOM2, PHDtm, TMAP, TMHMM2, TMPRED, and TOPPRED2, suggest helices with a length of 17−21 amino acids, while SPLIT4 and HMMTOP2 propose the helix with 30 and 25 amino acids, respectively.

For TM2, the shortest sequence was predicted by DAS with 14 amino acids followed by ALOM2, predicting a 17 residue helix. All other programs predict that the helix extends over 18−23 amino acids. TMAP is the only program which does not discriminate between the two helical regions and instead suggests a long helical stretch from Leu-8 to Leu-53. From the

* To whom correspondence should be addressed. Phone: +44(0)1865 275776. Fax: +44(0)1865 275234. E-mail: wolfgang.fischer@bioch.ox.ac.uk.
[†] Biomembrane Structure Unit, Department of Biochemistry.
[‡] Oxford Glycobiology Institute, Department of Biochemistry.
[§] Department of Physics.

10.1021/jm050721e CCC: $33.50
Published on Web 12/22/2005

A

```
NB        MNNATFNYTNVNPISHIRGSVIITICVSFTVILTVFGYIAKIFTKNNCTNNDIGLRERIK 60
VPU       -----------QPIPIVA---IVALVVAIIIAIVVWSIVIIEYRK-------I-LRQRKI 38
M2        ---MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGILHLILWI----------LDRLF 47
p7        ----------------------ALENLVILNAASLAGTHGLVSFL------------VF 25

NB        CSGCEPFCNKRDDISSPRTGVDIPSFILPGLNLSESTPN------------ 99
VPU       DRLIDRLIERAEDSGNESEG-EISALVEMGVEMGHHAPWDVDDL-------- 81
M2        FKCIYRFFKHGLKRGPSTEG--VPESMREEYRKEQQNAVDADDSHFVSIELE 97
p7        F-CFAWYLKGRWVPGAVYALYGMWPLLLLLLALPQRAYA------------ 63
```

B

```
                  10         20         30         40         50         60
                  |          |          |          |          |          |
          ALENLVILNA ASLAGTHGLV SFLVFFCFAW YLKGRWVPGA VYAFYGMWPL LLLLLALPQR AYA
ALOM2                    ******** *********           ********** *******
DAS                         *** **********              ******* *******
HMMTOP2          ***** ********** **********          ** ********** ********
MEMSAT1.5                                         ***** ********** *******
PHDtm                     ****** ********** ****     ** ********** ******
SPLIT4        ******** ********** ********** **       ** ********** ********** *
TMAP               *** ********** ********** ********** ********** ***
TMFINDER                    *** **********           * ********** ********
TMHMM2                   ******** ********** **    ***** ********** ********
TMPRED                      ***** ********** **    ***** ********** ******
TOPPRED2                 ********* ********** **      *** ********** ********
Consensus -----????? ??hhhhhHHH HHHHHHHHHH hh??-???hH HHHHHHHHHH HHHHHHh?-- ---
```

Figure 1. (A) Sequence alignment of individual strains of NB, from influenza B (SwissProt Entry Code P67908, VNB_INBLN), Vpu from HIV-1 (P05919, VPU_HV1H2), M2 from influenza A (P03491, VMT2_IABAN), and, e.g., p7 (P27958, POLG_HCVH). (B) Results from the run of multiple secondary prediction programs for the p7 sequence. Asterisks indicate the residues predicted to be within a helix. The consensus prediction is given at the bottom: H, consensus prediction; h, 6−8 methods predict helix; ?, less than 6 methods; - , no consensus.

above results a consensus prediction can be calculated according to a simple majority vote type procedure (Figure 1B). If more than eight methods predict a residue as helical, it is assigned an "H" in the consensus prediction. If between six and eight methods predict a helix residue then an "h" is assigned. If less than six methods predict a helix residue then a "?" is assigned. The regions that are strongly predicted as TM helices, i.e., assigned "H" or "h" in the consensus prediction, range from residues 13−32 for TM1 and from residues 39−57 for TM2.

Taking into account the results from the individual methods as well as the consensus prediction and the fact that the average length of an α-helix, as shown in a recent analysis of TM helices from known membrane protein structures,[23] is about 20 amino acids (see, e.g., refs 24 and 25 for other viral channel-forming proteins) the helical parts of p7 are proposed, for the purposes of this study, to include 23 amino acids ranging from Ala-10 to Leu-32 in TM1 and from Trp-36 to Pro-58 in TM2. This leaves a short three amino acid connecting loop: Lys-33, Gly-34, and Arg-35. The amphipathic TM1 helix includes residues Ser-12 and Ser-21, Thr-16, and His-17. Four phenylalanines are also present at positions 22, 25, 26, and 28. TM2 is largely highly hydrophobic with a high leucine content.

X-PLOR TM1/TM2 Monomers. On the basis of the above definitions for TM1 and TM2, two idealized antiparallel α-helices, separated by 9.4 Å,[26] were generated using the program X-PLOR. The optimal packing was determined by employing an interhelical configuration search protocol (SP) based on X-PLOR, very similar to the global search method.[27] Both methods screen the helix intermolecular energy landscape as this depends on the translational and rotational degrees of freedom of the two molecules.

According to the interhelical configuration SP, a number of models, with the TM1 and TM2 helices at different relative orientations, are generated. Two distinct ones have been built, each one starting from a different initial configuration of the TM1 and TM2 helices. In the first search protocol (SP1), the TM1 segment was placed in such a way that His-17 points away from the helix−helix interface. In the second search protocol (SP2), the relative orientation of the two helices is such that Gly-15 and Gly-18 from TM1 and Gly-39 and Gly-46 residues from TM2 all face the helix−helix interface.[5] In Figure 2, parts a (SP1) and b (SP2), the monomers are shown after the insertion of the small loop comprising the residues Lys-33, Gly-34, and Arg-35 (see below).

In the TM1/TM2 monomer generated from SP1 (Figure 2a), the TM2 helix axis forms an angle of 5° with respect to the TM1 axis, and it is also slightly kinked at Gly-39. The polar residues Ser-12, Thr-16, Ser-21, and Cys-27 are all close to the side of the helical interface. The glycines of both helices are placed at the helical interface. Furthermore, Phe-22, Phe-25, and Phe-26 of TM1 are located between Gly-39 and Gly-46 of TM2 while Leu-40 of TM2 is located between Gly-15 and Gly-18 of TM1, indicative for their role in helix−helix association.[5]

In the TM1/TM2 monomer, generated from SP2 (Figure 2b), the TM2 helix axis lies at an angle of 5° with respect to the TM1 at a distance of 9.0 Å. This structure is similar to the monomeric structure resulting from the first X-PLOR search (Figure 2a): the glycines of the two helices, the phenylalanines

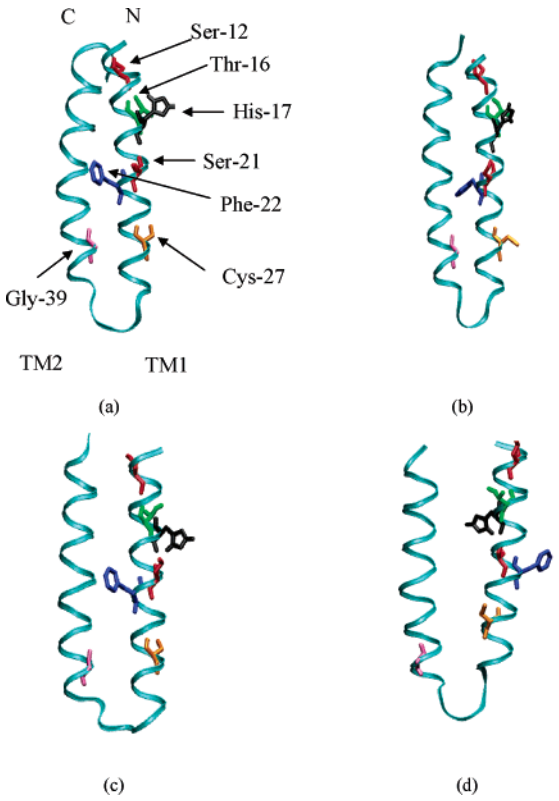

Figure 2. Models of p7 TM1/TM2 helical monomers with optimized interhelical packing which result from the different energetic searches: (a) from X-PLOR TM1/TM2 SP1 with the His-17 pointing away from the helical interface; (b) from X-PLOR TM1/TM2 SP2 with the Gly-15 and Gly-18 from TM1 and Gly-39 and Gly-46 from the TM2 fixed at the interhelical interface; (c) from docking with DOT; (d) from docking with GRAMM. In each model the linking residues have been added after the search/docking approach, and the resulting structure was then further energy minimized. Key residues are indicated by an arrow and named.

of TM1, and the leucines of TM2 are all facing the interface because of the initial constraints imposed.

DOT TM1/TM2 Monomers. With the use of the docking software DOT, one of the helices, TM1, is held fixed and TM2 is allowed to be rotated and translated as a rigid body after being aligned antiparallel. The matches are ranked according to their calculated exclusion volume, a parameter similar to the surface complementarity score, and the TM1/TM2 monomer with the highest value for its exclusion volume and the shortest distance between the TM2 N-terminus and TM2 C-terminus was selected.

The resulting TM1/TM2 monomer (Figure 2c) is strikingly similar to the one generated by the X-PLOR SP1 and SP2 (Figure 2, parts a and b). His-17 points away from the helical interface and so do all the other polar residues: Ser-12, Thr-16, Ser-21, and Cys-27. In addition, Phe-22, Phe-25, and Phe-26 of TM1 lie between Gly-39 and Gly-46 of TM2 and Leu-40 of TM2 between Gly-15 and Gly-18.

GRAMM TM1/TM2 Monomers. Of the 25 structures, both for the TM1 and TM2 helices derived from the X-PLOR protocol, 1 for each helix was used as starting structure for docking simulations with GRAMM. The program was run in the helix mode, i.e., the search was limited to helix pairs in antiparallel orientation, discarding configurations with large displacements along the helix axes and crossing angles larger than 10°.

In the GRAMM TM1/TM2 monomer (Figure 2d), the His-17 faces the helical interface and the other polar residues, Ser-21, Thr-16, and Cys-27, are pointing toward the intersubunit interface of the channel except Ser-13 which points to the channel pore.

Table 1. TM1/TM2 Monomers Resulted from the X-PLOR Search Protocol and the Docking Intermolecular Search[a]

TM1/TM2 monomer	energy (kcal/mol)	crossing angle (deg)	buried surface area (Å^2)
X-PLOR TM1/TM2 monomer SP1	−1409.3	5	998.1
X-PLOR TM1/TM2 monomer SP2	−1435.1	5	1037.1
DOT monomer	−1378.1	1.9	724.5
GRAMM monomer	−1350.0	4.3	527.9

[a] The sum of the electrostatic and van der Waals energies (kcal/mol) for all models after a final minimization as well as the crossing angle and the buried surface area between the two helices is given.

Table 2. p7 Oligomers Built by the TM1/TM2 Monomers Resulted from the Application of the X-PLOR Intermolecular Search Protocol and the DOT and GRAMM Software[a]

	X-PLOR SP1	X-PLOR SP2	DOT	GRAMM
tetramer	−5802.1	−5864.1	−5827.9	−5787.1
pentamer	−7283.2	−7419.2	−7380.5	−7354.2
hexamer	−8925.0	−361.0	−8784.7	−8851.7

[a] The values list the sum of the electrostatic and van der Waals energies (kcal/mol) after a final energy minimization.

Loop Insertion. The small loop (Lys-33, Gly-34, Arg-35) which links the TM1 and TM2 is also modeled by generating 10 loop structures with the SA/MD protocol. The first out of 10 structures is selected and attached to each of the above TM1/TM2 dimer models using the package Insight II. The resulting monomers are then subjected to 500 cycles of conjugate gradient minimization and are shown in Figure 2. The energy values, the crossing angles, and the buried surface-accessible areas are listed in Table 1. The TM1/TM2 monomer generated from X-PLOR SP2 has the lowest energy of about −1435.1 kcal/mol which correlates with the largest buried surface area of 1037.1 Å^2 and large crossing angle (5°). The other X-PLOR model has slightly higher energy and smaller surface area and crossing angle. The models derived from DOT and GRAMM adopt higher energies of about −1378 and −1350 kcal/mol, respectively. The surfaces areas are smaller for the X-PLOR models. However, the model derived from DOT is slightly larger (724.5 Å^2) than for GRAMM (527.9 Å^2). The crossing angles of the two models have been found to be approximately 2° and 4°.

Homooligomers. All of the TM1/TM2 monomers listed in Table 1, with and without the linking loop, are replicated and rotated by 90°, 72°, and 60° to form tetrameric, pentameric, and hexameric bundles of p7, respectively. These oligomers were subjected to 5000 cycles of conjugate gradient minimization with X-PLOR, and the resulting sums of the electrostatic and van der Waals energies for each of the bundles are listed in Table 2.

For the oligomers based on the X-PLOR TM1/TM2 monomer SP1 and GRAMM software the energies indicate an increase in stability with increasing oligomerization state in steps of approximately 1.200−1500 kcal/mol (Table 2). The minimum energy structures are found for the hexameric models X-PLOR TM1/TM2 monomer SP1 (−8925.0 kcal/mol) and GRAMM (−8851.7 kcal/mol). Hexameric models from X-PLOR TM1/TM2 monomer SP2 and DOT show higher energies of −361.0 kcal/mol, due to very high van der Waals interaction energy, and −8784.7 kcal/mol, respectively.

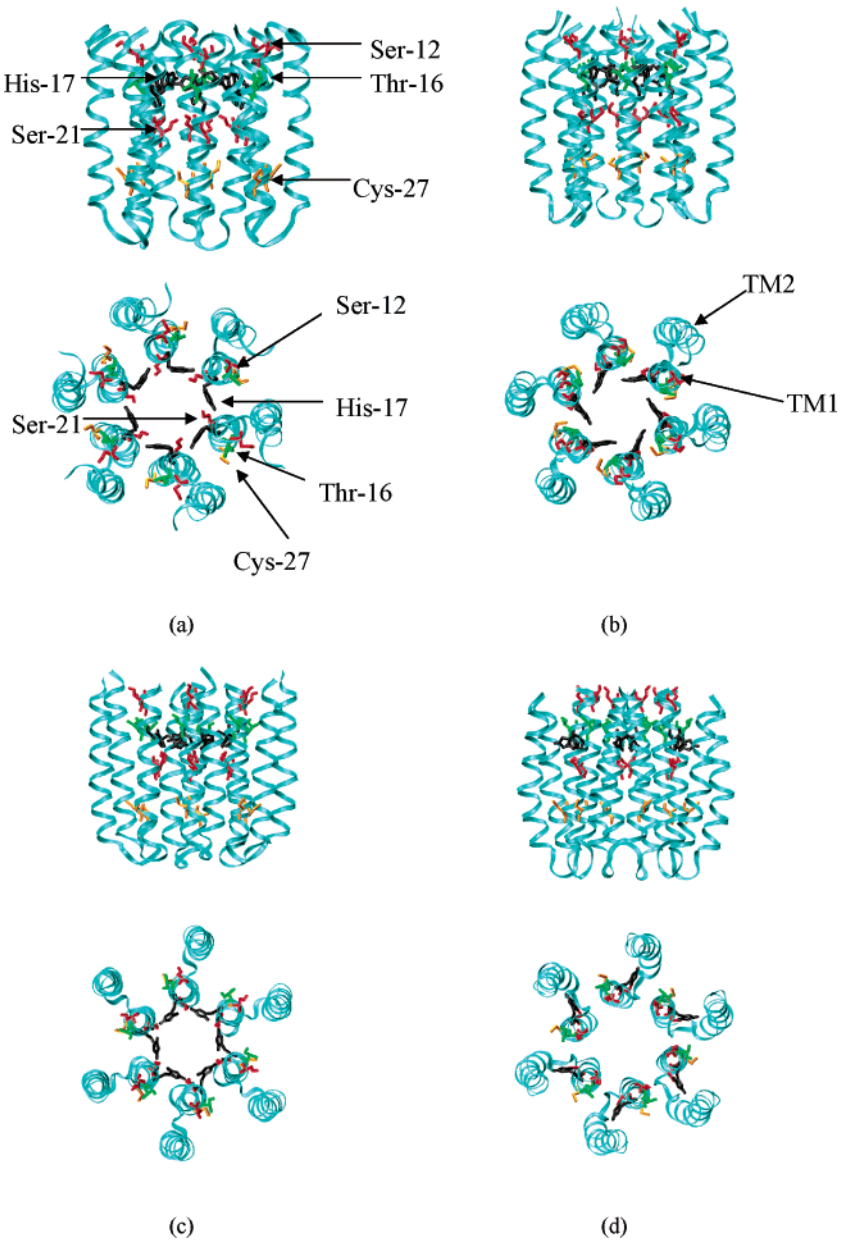

Figure 3. Top and side view of models of full p7 hexamers generated by the replication of the TM1/TM2 helix monomers shown in Figure 2: (a) XPLOR TM1/TM2 monomer SP1, (b) X-PLOR TM1/TM2 monomer SP2, (c) DOT, and (d) GRAMM. The structures resulting after an energy minimization are shown. The top view is from the N- and C-terminal side down to the extracellular linker.

The final structures of the p7 hexamers of each of the linked TM1/TM2 monomers are illustrated in Figure 3. The smaller bundles, i.e., tetramers and pentamers, show similar orientation of the respective amino acids and are, therefore, omitted in the figures. In the hexamer model constructed by the replication of the X-PLOR TM1/TM2 monomer SP1 (Figure 3a), His-17 and Ser-21 prominently point toward the center of the channel. The other polar residues, Ser-12 and Thr-16, are located at the protein−lipid interface, and they may be involved in intersubunit stabilization via hydrogen bond formation. Cys-27 is positioned in such a way that it cannot form disulfide bridges with the cysteines of the adjacent TM1 helices. There are two "aromatic belts" toward the protein−lipid interface in all of the models: one is formed by Trp-48, toward the helix ends, the other is formed by Trp-30 and Trp-36, toward the side with the linker (Figure 4, upper panel). Phe-25 and Phe-28 appear to be within the pore toward the side with the linker (Figure 4, lower panel). The other phenylalanines either face the lipid−protein interface (Phe-44) or are in close contact with TM2 (Phe-22 and Phe-26). The oligomers generated from X-PLOR TM1/TM2 monomer SP2 (Figure 3b) are structurally similar to the oligomers from the X-PLOR TM1/TM2 monomer SP1 (Figure 3a), because the glycine residues of the two helices are constrained at the helical interface. The oligomers differ only in the orientation of the constituent monomers: the TM1 helix is rotated ca. 20° clockwise compared to the model in Figure 3a.

The oligomers generated from the DOT TM1/TM2 monomer (Figure 3c) look almost the same as the oligomers resulting from X-PLOR TM1/TM2 monomer SP1 (Figure 3a) and X-PLOR TM1/TM2 monomer SP2 (Figure 3b) with respect to the orientation of the polar residues and the other structural features described above. In both of the oligomers of these models there is a ring of serines, Ser-21, located around the center of the channel.

In the oligomers generated from the GRAMM TM1/TM2 (Figure 3d) Ser-12 faces the pore. The overall packing of the bundle is different from one another because of the differences in the TM1/TM2 monomer buried surface area as mentioned

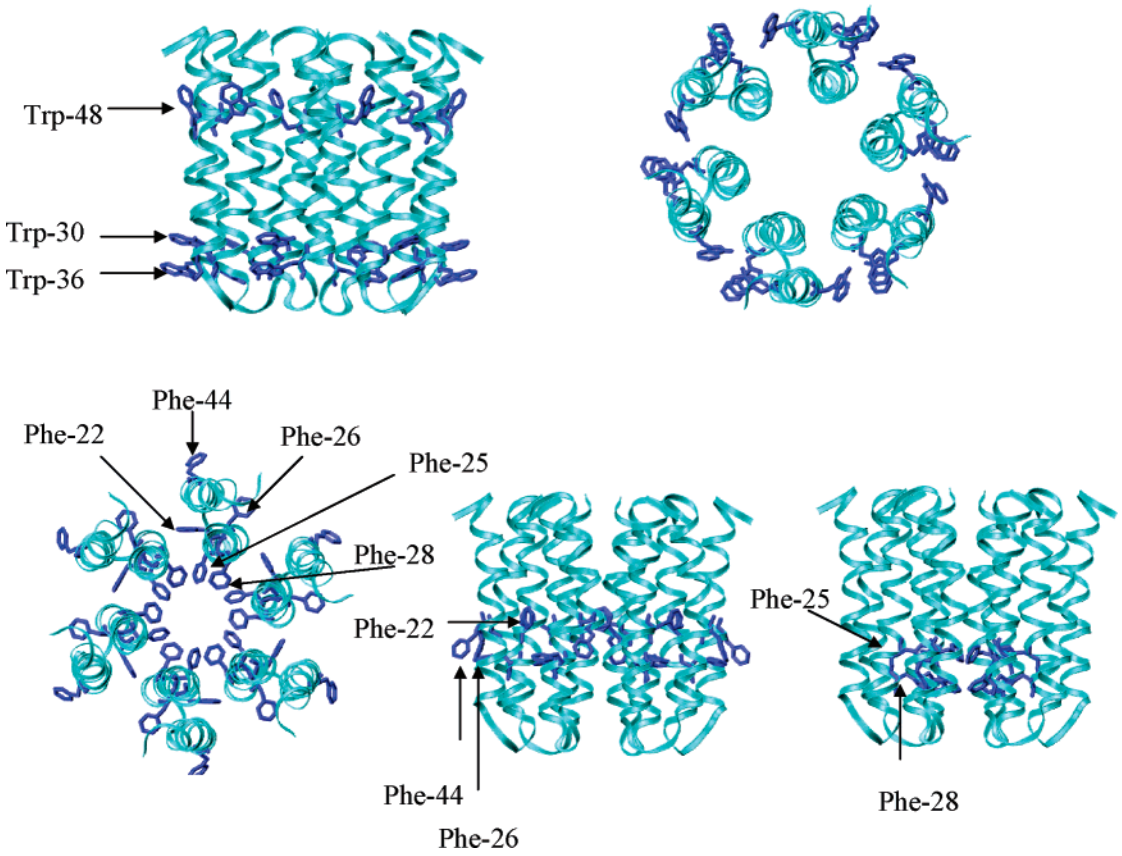

Figure 4. Upper panel: Side and top view of the p7 hexameric model resulting from X-PLOR TM1/TM2 monomer SP1. The tryptophan residues defining the channel's two aromatic belts, upper and lower, are denoted with a black arrow. Lower panel: The same model with the phenylalanine residues highlighted.

above. However, in these oligomers His-17 is closer the TM1 and TM2 interface.

The approximate dimensions of all the above models are ~1.5 nm for the inner and ~4.5 nm for the outer pore diameter which are slightly smaller than the values reported from TEM studies.[4]

As a result, the most plausible model proposed emerges from the hexameric assemblies of especially X-PLOR TM1/TM2 monomer SP1. The GRAMM model albeit with a low energy may not be considered further on the basis of the orientation of the histidines. The XPLOR TM1/TM2 monomer SP2 and DOT models do have higher interaction energy values and will currently not be considered further as well.

To assess whether the model X-PLOR TM1/TM2 monomer SP1 is reasonable a docking simulation method (AUTODOCK3.0) was used with amantadine in its protonated state ($pK_a = 9.0$,[28] Figure 5). Amantadine resides within the pore in the vicinity of His-17, interacting via hydrogen bonding with Ser-21. This position is independent of the protonation state of His-17. The calculated binding free energy is −4.32 kcal/mol with an estimated binding constant of $K_i = 68\ \mu M$.

Discussion

In this paper, the structure of the HCV ion channel p7 is modeled by computational methods. Given the lack of any direct experimental evidence related to the secondary or tertiary structure of the channel, the task set is similar to the protein-folding problem[22] and proceeds in a hierarchical manner: first the helical character of the transmembrane parts of the protein is assumed on the basis of findings from other viral ion channels,[29] then multiple secondary prediction methods are performed to determine the extent of each TM helix. Subsequently, the packing of TM1 and TM2 helices is investigated by a generation of a number of models either by a configuration search protocol or by a helix−helix docking approach. The prediction strategy embodies the two-stage model of membrane protein folding[30] which says that in the first stage independently stable TM helices are formed across the bilayer and in the second stage the helices interact with each other to form the final folded monomeric protein structure. In the case of p7, it is assumed that a possible following stage leads to the final oligomerization of the monomer.

The two docking methods employed, DOT and GRAMM, have been mainly developed for protein docking of globular proteins and used here for membrane proteins. It is surprising that the energy values for the TM1/TM2 monomers are only slightly higher than those from the coarse grained search XPLOR models. Even though the DOT TM1/TM2 monomer seems to have a plausible orientation of the residues the energy values for the bundle are high. In contrast, GRAMM delivers a nonphysiologically relevant orientation, but in a hexameric bundle, the energy is quite low. Thus, at the current stage, these programs need to be used with caution for studying helix−helix interactions of the TM parts of membrane proteins.

Previous secondary structure prediction results for p7[5] propose that the first helix, TM1, is only 14 residues long (amino acids 19−32), which would be too short to span the lipid bilayer. The authors propose that it may extend up to position 13 or even 10 as indicated by other prediction methods, an assumption that is supported by the results presented here (Figure 1). The analysis of Carrère-Kremer et al.[5] and the present analysis both suggest that the TM2 helix extends from residues 36 to 58.

The favored p7 hexameric model XPLOR SP1 (Figure 3a) displays the maximum number of polar residues facing the pore of the channel. The central ring formed by the Ser-21 residues

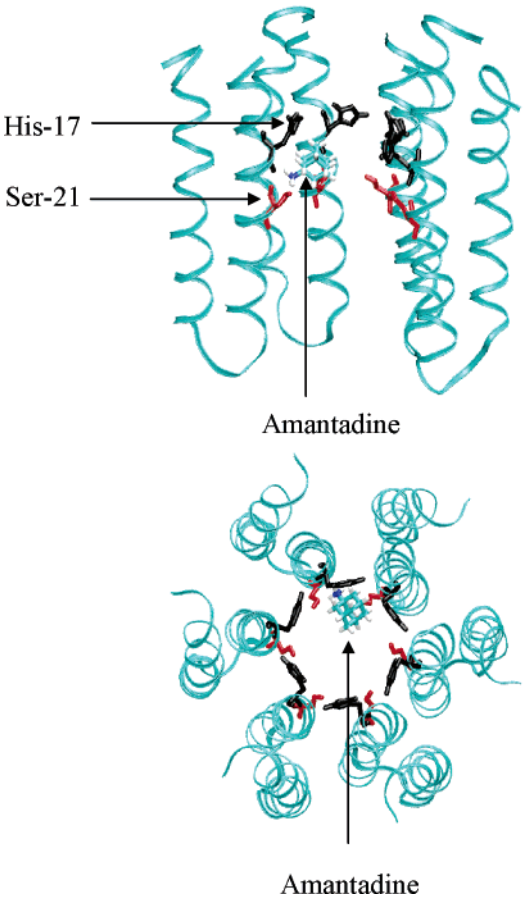

Figure 5. Side (top) and top view (bottom) of amantadine (CPK color code) within p7 model X-PLOR TM1/TM2 monomer SP1. His-17 and Ser-21 are shown in black and red, respectively. To view the location of amantadine within the bundle (side view) two monomers have been removed for clarity. The top view is from the N- and C-terminal side down to the extracellular linker.

at the center of the channel is reminiscent of similar serine rings observed in the acetylcholine receptor.[31] A particularly interesting feature of these models is the orientation of His-17 toward the interior of the pore, also observed in the M2 proton channel where His-37 plays a key role in pH detection, gating, and proton conduction.[32,33] In p7, the histidine is located in position 17 (only two residues away from position 19, which is designated as the start of TM1 helix by most secondary structure prediction methods) and is likely to be part of the channel's inner polar TM1 helix. Amantadine blocks the M2 channel by binding to the pore-lining His-37 residue;[19,32,34] amantadine has also been found to block p7,[4] supporting the conjecture that the His-17 residues line the pore. Also, the preliminary docking results in this study suggest amantadine to reside in the vicinity of the His-17. A possible way to further elucidate the role of His-17, through computational means, is to perform molecular dynamics simulations of the TM1 helix with various lengths embedded in a solvated lipid bilayer and examine whether this part of the helix, containing His-17, remains stable. This approach has been successfully applied for the definition transmembrane helix of M2[20] and other viral ion-channel-forming proteins.[35] It is possible that His-17 plays a role similar to that of His-37 of M2 in gating and suggests also that p7 may need to be tested upon conductance of protons. In M2 Trp-41 is one turn of the helix apart from His-37 and possibly involved in gating.[36−39] For p7, His-17 and Phe-25 are two turns of a helix apart and may have a similar role as that of Trp-41 in M2.

In the models presented in this paper, the other aromatic residues such as tryptophans and phenylalanines are placed so that their position enables them to anchor the protein within the lipid bilayer. The ring of tryptophans is also seen in other membrane proteins such as porins[40,41] and bacteriorhodopsin.[42,43]

Conclusions

A pure computational hierarchical methodology[44] has been applied for the modeling of a homooligomeric assembly of the channel-forming protein p7 from HCV in the absence of any high-resolution structural and any other biochemical data regarding the TM stretches. The results suggest a hexameric assembly. The coarse grained search of the orientational space proposes the most plausible model in respect of its structure and energy. The model suggests that histidines are involved in ion channel gating.

Materials and Methods

Sequence Alignment. Sequences were downloaded from the SwissProt database and copied into the EBI-ClustalW server (www.ebi.ac.uk/clustalw/).

Secondary Structure Prediction. The sequence of the p7 peptide, which was used for experimental studies,[3]

ALENLVILNA[10] ASLAGTHGLV[20] SFLVFFCFAW[30] YLKGR-WVPGA[40] VYAFYGMWPL[50] LLLLALPQR[60] AYA

was taken, and 11 secondary structure prediction methods (ALOM2, DAS, HMMTOP 2.0, MEMSAT1.5, PHDhtm, SPLIT-4.0, TMAP, TM-Finder, TMHMM2.0, TMPred, TopPred2) available on a local bioinformatics server[23] were employed to determine the start and end points of the two proposed helical regions of p7. However, this does not represent the whole range of prediction methods. It leaves out others such as, e.g., kPROT (http://bioinfo.weizmann.ac.il/kPROT/). All of the above methods utilize different criteria in order to determine a part of a protein sequence as transmembrane. For example, HMMTOP2.0[45] and TMHMM2.0[46] make use of hidden Markov models, PHDhtm[47] is based on a neural network method in which information from multiple aligned sequences is included; MPEX uses hydropathy plots in combination with thermodynamics principles.[48] Given this, it is preferable to extract a certain consensus answer from the results of multiple secondary structure prediction methods rather than depending on the results of only one of them.[23]

SA/MD Protocol. Tetrameric, pentameric, and hexameric bundles of each of the predicted TM helices, TM1 and TM2, were generated using a simulated annealing/molecular dynamics (SA/MD) protocol based on the program X-PLOR.[49] The procedure has been extensively described elsewhere.[50] In brief, the SA/MD protocol consists of two stages: In stage 1, the bundles were constructed with idealized α-helices based on the positions of the Cα atoms of the peptide unit. The initial tilt angle of the helices was set at 5°. All other atoms of the individual amino acid side chains were superimposed on the respective Cα atoms. During stage 1, these side chain atoms "evolve" from the Cα atoms, which were held restrained in their original position. By beginning the annealing at 1000 K, weights for bond length, bond angle planarity, and chirality were gradually increased. A repulsive van der Waals term was introduced after an initial delay. The bundle was then cooled to 300 K, in steps of 10 K per 0.5 ps. van der Waals radii were reduced to 80% of the original value to allow atoms to pass each other. The annealing was repeated five times, and five structures were obtained for each bundle. Each structure from stage 1 was used for five molecular dynamics runs in stage 2.

In stage 2 the initial velocities corresponded to 500 K. Harmonic restraints initially held the Cα atoms but were released when the temperature was decreased from 500 to 300 K. At this stage distance restraints were also introduced. At 300 K a constant temperature molecular dynamics simulation for 5 ps was performed, followed by 1000 steps of conjugate gradient minimization. Then the Cα atoms were also allowed to move. In stage 2 electrostatic interac-

tions were introduced into the potential energy function. The main chain atoms obtained their charges corresponding to the PARAM19 parameter set. Partial charges on the side chain atoms of polar side chains were gradually scaled up from 0.005 to 0.4 times their full value during the temperature reduction from 500 to 300 K. The scaling factor 0.4 was also applied during the 5 ps dynamics and energy minimization. A distance-dependent dielectric function was used with a switching function smoothly truncating the distant electrostatic interactions. In total, 25 structures were obtained in stage 2. For the molecular dynamics in this protocol the CHARMM force field[44] and the Verlet integration method were used.

X-PLOR Interhelical Search Protocols. Two distinct interhelical search protocols (SP) were performed with X-PLOR. The model generation was performed in the following way: In the first protocol, SP1, the TM1 was kept fixed and TM2 allowed moving in three rotational degrees of freedom: around the *z*-axis, φ_1, around the *y*-axis, φ_2, and around its helical axis, φ_3. For the angle φ_1 three values were used, −45°, 0°, and 45°; for the angle φ_2 five values were used, −10°, −5°, 0°, 5°, and 10°, and for the angle φ_3, 0°, 60°, 120°, 180°, 240°, and 300° are used. The combination of all the above rotations yielded a total of 90 TM1/TM2 dimers. In the second protocol, SP2, TM2 was placed at different distances from TM1, at 8.0, 8.5, 9.0, 9.5, and 10.0 Å, and it was rotated around the *y*-axis at −10°, −5°, 0°, 5°, and 10° yielding a total of 25 TM1/TM2 dimers. The value for the angle φ_1 in SP2 was chosen so as to bring the conserved Gly-15 and Gly-17 of TM1 and Gly-39 and Gly-46 of TM2 facing each other at the interhelical interface. This is because glycine residues are often observed in transmembrane α-helices oligomerization.

Each of the 175 models were put through a single energy minimization step using the X-PLOR protocol as described earlier (cycle of the SA/MD protocol). The energy calculation for each model includes bond length, bond angle, proper dihedral, improper dihedral, and van der Waals and electrostatic energy terms. The model with the lowest sum of electrostatic and van der Waals energy value is shown in Figure 2. Finally, the short linking loop was attached to the C- and N-terminal ends of TM1 TM2 to obtain the TM1/TM2 monomer.

Helix−Helix Docking. Two academically available programs were used: DOT[51] (http://www.sdsc.edu/CCMS/DOT/) and GRAMM[21] (http://www.bioinformatics.ku.edu/research/vakser/resources/gramm/gramm1/) to dock the TM1 and TM2 helices.

A systematic search over the six degrees of freedom of the two helices was performed using DOT. For each helical complex, the sum of a Poisson−Boltzmann electrostatic energy of interaction and van der Waals energy, each represented as a grid-based correlation function, were computed. The program evaluated the energy of interaction for many orientations of the moving molecule, in this experiment the TM2 helix, and scored the resulting structures by the electrostatic energy, the van der Waals energy, or the composite sum of both. The potential grid for the fixed molecule, TM1, was calculated using the program UHBD.[52,53] The docking parameters were set to their default values. The lowest and the highest value allowed for the still-molecule mask grid were −1.0 and 1000, respectively. The lowest and the highest value for the potential mask grid were −6.0 and 6.0, respectively. The grid step was 1.0 Å, and the grid size was 128.

An exhaustive 6-dimensional search through the relative rotations and translations of the two helices was performed using GRAMM yielding a number of different TM1/TM2 complexes ranked according to a surface complementarity score.[54] The values of the other docking parameters were the ones suggested for high-resolution docking.[55] The grid spacing[54,56,57] was 1.7 Å, and the repulsion part of the potential[55] function 30.0 (in arbitrary units). The projection of an atom was "yes−no", as spheres with the van der Waals radius. The attraction double range[54,58] was set to zero. The potential of all atoms was projected as spheres with the corresponding van der Waals radius. The angle for the rotations was 10°, and the first 1000 matches were kept as outputs. The match structure with the highest surface complementarity score was selected.

The buried surface area for a given TM1/TM2 monomer was calculated as the difference of the surface-accessible area of the whole TM1/TM2 monomer and the surface-accessible area of each of the TM1 and TM2 helices using the NACCESS program.

Ligand Docking. Amantadine in its protonated state was generated using WebLab ViewerPro. For the docking simulation the AUTODOCK3.0 software was used. The docking behavior was predicted by a linear regression analysis in an AMBER force field and Monte Carlo simulated annealing using a Lamarckian Genetic Algorithm.[59] A population size of 50, with 30 500 000 maximum number of energy evaluations, 9 700 000 maximum number of generations, gene mutation rate of 0.02, and cross over of 0.8 were used for the global search algorithm. The number of iterations for the local search is 300 using the Solis and Wets algorithm.[60] The docking run contained 50 runs to generate clusters of ultimate docking positions (rmsd tolerance 1 Å).

All the figures presented in this paper were produced with the VMD.[61]

Acknowledgment. W.B.F. thanks the Bionanotechnology IRC for financial support. This work was supported by an Oxford Glycobiology Institute endowment Grant all by United Therapeutics Inc. Thanks to Professor Anthony Watts (Oxford) for providing laboratory space and C. G. Kim (Oxford) for helpful discussions about the docking simulation.

References

(1) Lohmann, V.; Korner, F.; Koch, J.; Herian, U.; Theilmann, L.; et al. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **1999**, *285*, 110−113.
(2) Premkumar, A.; Wilson, L.; Ewart, G. D.; Gage, P. W. Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride. *FEBS Lett.* **2004**, *557*, 99−103.
(3) Pavlovic, D.; Neville, D. C.; Argaud, O.; Blumberg, B.; Dwek, R.; et al. The hepatitis C virus protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6104−6108.
(4) Griffin, S. D. C.; Beals, L. P.; Clarke, D. S.; Worsfold, O.; Evans, S. D.; et al. The p7 protein of the hepatitis C virus forms a virus channel that is blocked by the antiviral drug, amantadine. *FEBS Lett.* **2003**, *535*, 34−38.
(5) Carrere-Kremer, S.; Montpellier-Pala, C.; Cocquerel, L.; Wychowski, C.; Penin, F.; et al. Subcellular localisation and toplogy of the p7 polypeptide of the hepatitis C virus. *J. Virol.* **2002**, *76*, 3720−3730.
(6) Griffin, S. D. C.; Harvey, R.; Clarke, D. S.; Barkley, W. S.; Harris, M.; et al. A conserved loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity but is dispensable for localisation to mitochondria. *J. Gen. Virol.* **2004**, *85*, 451−461.
(7) Kovacs, F. A.; Cross, T. A. Transmembrane four-helix bundle of influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophys. J.* **1997**, *73*, 2511−2517.
(8) Marassi, F. M.; Ma, C.; Gratkowski, H.; Straus, S. K.; Strebel, K.; et al. Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14336−14341.
(9) Wray, V.; Kinder, R.; Federau, T.; Henklein, P.; Bechinger, B.; et al. Solution structure and orientation of the transmembrane anchor domain of the HIV-1-encoded virus protein U by high resolution and solid-state NMR spectroscopy. *Biochemistry* **1999**, *38*, 5272−5282.
(10) Kukol, A.; Adams, P. D.; Rice, L. M.; Brunger, A. T.; Arkin, I. T. Experimentally based orientational refinement of membrane protein models: a structure for the influenza A M2 H^+ channel. *J. Mol. Biol.* **1999**, *286*, 951−962.
(11) Kukol, A.; Arkin, I. T. Vpu transmembrane peptide structure obtained by site-specific Fourier transform infrared dichroism and global molecular dynamics searching. *Biophys. J.* **1999**, *77*, 1594−1601.
(12) Fischer, W. B.; Pitkeathly, M.; Wallace, B. A.; Forrest, L. R.; Smith, G. R.; et al. Transmembrane peptide NB of influenza B: a simulation, structure, and conductance study. *Biochemistry* **2000**, *39*, 12708−12716.
(13) Adams, P. D.; Engelman, D. M.; Brunger, A. T. Improved prediction of the structure of the dimeric transmemebrane domain of Glycophorin A obtained through global searching. *Proteins* **1996**, *26*, 257−261.
(14) Gottschalk, K. E.; Adams, P. D.; Brunger, A. T.; Kessler, H. Transmembrane signal transduction of the alpha(IIb)beta(3) integrin. *Protein Sci.* **2002**, *11*, 647−665.

(15) Kochva, U.; Leonov, H.; Arkin, I. T. Modeling the structure of the respiratory syncytial virus small hydrophobic protein by silent mutation analysis of global searching molecular dynamics. *Protein Sci.* **2003**, *12*, 2668−2674.
(16) Torres, J.; Briggs, J. G.; Arkin, I. T. Convergence of experimental, computational and evolutionary methods predicts the presence of a tetrameric form for CD3-ζ. *J. Mol. Biol.* **2002**, *316*, 375−384.
(17) Torres, J.; Briggs, J. G.; Arkin, I. T. Contribution of energy values to the analysis of global searching moleulcar dynamics simulations of transmembrane helical bundles. *Biophys. J.* **2002**, *82*, 3063−3071.
(18) Choma, C. T.; Tieleman, D. P.; Cregut, D.; Serano, L.; Berendsen, H. J. C. Towards the design and computational characterisation of a membrane protein. *J. Mol. Graphics Modell.* **2001**, *20*, 219−234.
(19) Duff, K. C.; Gilchrist, P. J.; Saxena, A. M.; Bradshaw, J. P. Neutron diffraction reveals reveals the site of amantadine blockaded in influenza A M2 ion channel. *Virology* **1994**, *202*, 287−293.
(20) Forrest, L.; Tieleman, D. P.; Sansom, M. Defining the transmembrane helix of M2 protein from influenza A by molecular dynamics simulations in a lipid bilyer. *Biophys. J.* **1999**, *76*, 1886−1896.
(21) Vakser, I. A.; Jiang, S. Strategies for modelling the interactions of transmembrane helices of G protein-coupled receptors by geometric complemntarity using the GRAMM computer algorithm. *Methods Enzymol.* **2002**, *343*, 313−328.
(22) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins* **2002**, *47*, 409−443.
(23) Cuthbertson, J. M.; Doyle, D. A.; Sansom, M. S. Transmembrane helix prediction: A comparative evaluation and analysis. *Protein Eng. Des. Sel.* **2005**, *18*, 295−308.
(24) Ma, C.; Marassi, F. M.; Jones, D. H.; Straus, S. K.; Bour, S.; et al. Expression, purification, and activities of full-length and truncated versions of the integral membrane protein Vpu from HIV-1. *Protein Sci.* **2002**, *11*, 546−557.
(25) Pinto, L. H.; Dieckmann, G. R.; Gandhi, C. S.; Papworth, C. G.; Braman, J.; et al. A functionally defined model for the M2 proton channel of influenza A virus suggests a mechanism for its ion selectivity. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11301−11306.
(26) Chothia, C.; Levitt, M.; Richardson, D. Helix to helix packing in proteins. *J. Mol. Biol.* **1981**, *145*, 215−250.
(27) Adams, P. D.; Arkin, I. T.; Engelman, D. M.; Brunger, A. T. Computational searching amd mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat. Struct. Biol.* **1995**, *2*, 154−162.
(28) Spector, R. Transport of amantadine and remantadine through the blood−brain barrier. *Pharmacol. Ther*. **1988**, *244*, 516−519.
(29) Fischer, W. B.; Sansom, M. S. P. Viral ion channels: structure and function. *Biochim. Biophys. Acta* **2002**, *1561*, 27−45.
(30) Popot, J. L.; Engelman, D. M. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **1990**, *29*, 4037−4041.
(31) Miyazawa, A.; Fujiyoshi, Y.; Unwin, N. Structure and gating mechanism of the acetylcholine receptor. *Nature* **2003**, *423*, 949−955.
(32) Wang, C.; Lamb, R. A.; Pinto, L. H. Activation of M2 ion channel of influenzla A virus−A role for the transmembrane domain histidine residue. *Biophys. J.* **1995**, *69*, 1363−1371.
(33) Gandhi, C. S.; Shuck, K.; Lear, J. D.; Dieckmann, G. R.; DeGrado, W. F.; et al. Cu(II) inhibition of the proton translocation machinery of the influenza A virus M-2 protein. *J. Biol. Chem.* **1999**, *274*, 5474−5482.
(34) Salom, D.; Hill, B. R.; Lear, J. D.; DeGrado, W. F. pH-dependent tetramerisation and amantadine binding of the transmembrane helix of M2 from influenza A virus. *Biochemistry* **2000**, *39*, 14160−14170.
(35) Fischer, W. B.; Forrest, L. R.; Smith, G. R.; Sansom, M. S. P. Transmembrane domains of viral ion channel proteins: a molecular dynamics simulation study. *Biopolymers* **2000**, *53*, 529−538.
(36) Holsinger, L. J.; Nichani, D.; Pinto, L. H.; Lamb, R. A. Influenza A virus M2 ion channel protein: a structure−function analysis. *J. Virol.* **1994**, *68*, 1551−1563.
(37) Sansom, M. S. P.; Kerr, I. D.; Smith, G. R.; Son, H. S. The influenza A virus M2 channel: a molecular modelling and simulation study. *Virology* **1997**, *233*, 163−173.
(38) Shuck, K.; Lamb, R.; Pinto, L. H. Analysis of the pore structure of the influenza A virus M_2 ion channel by the substituted-cystein accessibility method. *J. Virol.* **2000**, *74*, 7755−7761.
(39) Czabotar, P. E.; Martin, S. R.; Hay, A. J. Studies of structural changes in the M2 proton channel of influenza A virus by tryptophan fluorescence. *Virus Res.* **2004**, *99*, 57−61.
(40) Weiss, M. S.; Kreusch, A.; Schiltz, E.; Nestel, U.; Welte, W.; et al. The structure of porin from rhodobacter capsulatus at 1.8 Å resolution. *FEBS Lett.* **1991**, *280*, 379−382.
(41) Pautsch, A.; Schulz, G. E. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **2000**, *298*, 273−282.
(42) Pebay-Peyroula, E.; Rummerl, G.; Rosenbusch, J. P.; Landau, E. M. X-ray structure of bacteriorhodopsin at 2.5 Angstroms from microcrystals grown in lipid cubic phases. *Science* **1997**, *277*, 1676−1681.
(43) Sass, H. J.; Büldt, G.; Gessenich, R.; Hehn, D.; Neff, D.; et al. Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin. *Nature* **2000**, *406*, 649−653.
(44) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; et al. CHARMM: A program for macromolecular energy, minimization, and molecular dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(45) Tusnady, G. E.; Simon, I. Principles governing amino acid composition of integral membrane proteins: Applications to topology prediction. *J. Mol. Biol.* **1998**, *283*, 489−506.
(46) Krogh, A. L. B.; von Heijne, G.; Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567−580.
(47) Rost, B. PHD: predicting 1D protein structure by profile based neural networks. *Methods Enzymol.* **1996**, *266*, 525−539.
(48) White, S. H.; Wimley, W. C. Membrane protein folding and stability: Physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319−365.
(49) Brunger, A. T. *X-PLOR, Version 3.1: A system for X-ray crystallography and NMR*; Yale University Press: New Haven, CT, 1993.
(50) Kerr, I. D.; Sansom, M. S. P. Hydrophilic surface maps of α-helical pore forming peptides. *Eur. Biophys. J.* **1993**, *22*, 269−277.
(51) Mandell, J. C.; Roberts, V. A.; Pique, M. E.; Kotlovyi, M.; Mitchell, J. C.; et al. Protein docking using continum electrostatics and geometric fit. *Protein Eng.* **2001**, *14*, 105−113.
(52) Davis, M. E.; Madura, J. D.; Lutty, B. A.; McCammon, J. A. Electrostatics and diffussion of molecules in solution: Simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **1991**, *62*, 187−197.
(53) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Lutty, B. A.; et al. Electrostatics and Diffusion in solution: Simulations with the University of Houston Brownian Dynamics program. *Comput. Phys. Commun.* **1995**, *91*, 57−95.
(54) Katchalski-Kazir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. A.; Aflalo, C.; et al. Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2195−2199.
(55) Vakser, I. A. Long-distance potentials: An approach to the multiple minima problem in ligand receptor interaction. *Protein Eng.* **1996**, *9*, 37−41.
(56) Vakser, I. A. Protein docking for low resolution structures. *Protein Eng.* **1995**, *8*, 371−377.
(57) Vakser, I. A. Low resolution docking: Prediction of complexes for underdetermined structures. *Biopolymers* **1996**, *39*, 455−464.
(58) Vakser, I. A.; Aflalo, C. Hydrophobic docking: a proposed enhancement to molecular recognition techniques. *Proteins* **1994**, *20*, 320−329.
(59) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(60) Solis, F. J.; Wets, R. J. Minimization by random search techniques. *Math. Oper. Res.* **1981**, *6*, 19−30.
(61) Humphrey, W.; Dalke, A.; Schulten, K. VMD−Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

JM050721E