The Transmembrane Oligomers of Coronavirus Protein E

Jaume Torres,* Jifeng Wang,* Krupakar Parthasarathy,* and Ding Xiang Liu*†
*School of Biological Sciences, Nanyang Technological University, Singapore; and †Institute of Molecular and Cell Biology, Proteos, Singapore

ABSTRACT We have tested the hypothesis that severe acute respiratory syndrome (SARS) coronavirus protein E (SCoVE) and its homologs in other coronaviruses associate through their putative transmembrane domain to form homooligomeric α-helical bundles in vivo. For this purpose, we have analyzed the results of molecular dynamics simulations where all possible conformational and aggregational space was systematically explored. Two main assumptions were considered; the first is that protein E contains one transmembrane α-helical domain, with its N- and C-termini located in opposite faces of the lipid bilayer. The second is that protein E forms the same type of transmembrane oligomer and with identical backbone structure in different coronaviruses. The models arising from the molecular dynamics simulations were tested for evolutionary conservation using 13 coronavirus protein E homologous sequences. It is extremely unlikely that if any of our assumptions were not correct we would find a persistent structure for all the sequences tested. We show that a low energy dimeric, trimeric and two pentameric models appear to be conserved through evolution, and are therefore likely to be present in vivo. In support of this, we have observed only dimeric, trimeric, and pentameric aggregates for the synthetic transmembrane domain of SARS protein E in SDS. The models obtained point to residues essential for protein E oligomerization in the life cycle of the SARS virus, specifically N15. In addition, these results strongly support a general model where transmembrane domains transiently adopt many aggregation states necessary for function.

INTRODUCTION

Coronaviruses, which belong to the family *Coronaviridae*, cause common colds in humans and are responsible for serious diseases in other species. Recently, one of its members has been found to be the causative agent of the severe acute respiratory syndrome (SARS) (Rota et al., 2003). Coronaviruses are surrounded by a lipid bilayer, or envelope, which typically embeds three proteins: spike (S), matrix (M), and the E protein. The envelope surrounds a nucleocapsid, containing the viral RNA and nucleocapsid (N) protein. Proteins S, M, and N have been studied for their important roles in receptor binding and virion budding. For example, the envelope spike protein S mediates attachment to cellular receptors and entry by fusion with cell membranes, whereas the matrix protein M is involved in budding and interacts with N and S proteins (Opstelten et al., 1995; Narayanan et al., 2000).

The significance of the E protein, however, has proved more elusive, but appears to be critical for viral budding, as charged-to-alanine mutations in mouse hepatitis virus (MHV) have been found to produce dramatic morphological changes in the virions (Fischer et al., 1998). Additionally, although in many coronaviruses expression of M protein on its own is not sufficient to produce virus-like particles, co-expression of proteins M and E can readily produce them (Bos et al., 1996; Vennema et al., 1996; Baudoux et al., 1998; Corse and Machamer, 2000). Proteins M and E have been found to interact via their cytoplasmic domains in pre-Golgi compartments (Lim and Liu, 2001). Another role suggested for protein E has been in promoting apoptosis (An et al., 1999; Chen et al., 2001), an effect that can be opposed by Bcl-2. Further, recent data (Liao et al., 2004) suggests that SARS coronavirus E protein (SCoVE) can increase membrane permeability and may have ion channel activity.

Despite, or because of, its small size, the topology of protein E is still a matter of controversy. Some reports (Corse and Machamer, 2000) have suggested that protein E in IBV traverses the Golgi lipid bilayer once, with the N-terminus facing the Golgi lumen and the C-terminus facing the cytoplasm. Another group (Maeda et al., 2001) has suggested that protein E in MHV traverses the lipid bilayer twice, whereby both N- and C-termini of the protein would reside in the cytoplasm, which is topologically equivalent to the interior of the viral envelope. Even more recently, based on in vitro biophysical studies (Arbely et al., 2004) a short hairpin (12 amino acids long) has been suggested for the putative transmembrane domain of SCoVE.

The explanation for these seemingly conflicting reports may be either of experimental origin or perhaps related to the protein's reported varied functionality. In any case, protein E clearly has the potential to perturb or permeabilize lipid bilayers (Fischer et al., 1998), but the structural determinants involved (pores, hairpins) have not been clearly defined.

To predict a possible transmembrane oligomer of protein E, we have worked under the assumption that protein E contains one transmembrane domain with its N- and C-termini in opposite sides of the membrane (Corse and Machamer, 2000). We have then performed global searching

Submitted August 22, 2004, and accepted for publication November 1, 2004.

Jaume Torres and Jifeng Wang contributed equally to this work.

Address reprint requests to Jaume Torres, Tel.: 65-6316-2857; Fax: 65-6791-3856; E-mail: jtorres@ntu.edu.sg.

0006-3495/05/02/1283/08 $2.00

doi: 10.1529/biophysj.104.051730

molecular dynamics simulations (Adams et al., 1995) using only the transmembrane sequence of protein E (TME). As the oligomeric size of the hypothetic bundle is not known, we explored different oligomeric sizes, from dimers to hexamers. This procedure was performed on 13 different sequence variants, to select a model that would be evolutionarily conserved (Briggs et al., 2001). The latter strategy has already been used successfully to predict correct models for transmembrane peptides known to form dimers (Briggs et al., 2001), trimers (Kukol et al., 2002), tetramers (Torres et al., 2002a,b), or pentamers (Torres et al., 2002a) for which experimental data was available a priori, and the validity of the predictions could be readily assessed. In this study in contrast, neither the precise topology of protein E, nor the oligomeric size, helix tilt, and helix rotational orientation of the hypothetic α-helical bundle were known.

We reasoned that if a structure could be found that was not affected by any conservative mutation, and if a large number sequences with rather low similarity were used, it would be extremely unlikely that this should happen by chance and therefore the structure should be present in vivo. Surprisingly, we found not one, but four homooligomeric models, a dimer, a trimer and two pentamers, for the transmembrane domain of the coronavirus E protein.

MATERIALS AND METHODS

Homologous sequences and predicted transmembrane domains

The sequence of SCoVE (Fig. 1) was obtained from Swiss-Prot and TrEMBL (http://ca.expasy.org/sprot/sprot-top.html), and its homologous sequences were obtained using the FASTA server (http://www.ebi.ac.uk/fasta33). In total, 13 homologous sequences were used in this study with a minimum similarity of 17% in their predicted transmembrane domain (see Fig. 2). The complete names of these sequences, the abbreviation used in Fig. 2 (within parentheses), and Swiss-Prot entries are: SARS coronavirus E protein (SCoVE), P59637; small envelope protein from SARS coronavirus BJ01 (SCoV_BJ01), Q6QJ39; envelope protein from feline coronavirus (FCoV), O12296; envelope protein from canine coronavirus (CCoV), Q7T6T0; envelope protein from canine enteric coronavirus, strain Insavc-1 (CCoV_Insacv1), VEMP_CVCAI; envelope protein from porcine transmissible gastroenteritis coronavirus, strain Purdue (TGEV_purdue), VEMP_CVPPU; envelope protein from porcine respiratory coronavirus, strain RM4 (PrCoV_RM4), VEMP_CVPRM; putative small membrane protein from porcine transmissible gastroenteritis coronavirus, strain FS772/70 (TGEV_FS772/70), VSMP_CVPFS; small membrane protein E from porcine hemagglutinating encephalomyelitis virus (PHEV), Q84730; small membrane protein from Rat sialodacryoadenitis coronavirus (RCoV), Q9IKC8; small membrane protein from Murine hepatitis virus (MHV), O72007; envelope protein from Human coronavirus, strain 229E (HcoV_229E), VEMP_CVH22; and putative small membrane protein from Avian infectious bronchitis virus, strain M41 (IBV_M41), VSMP_IBVM.

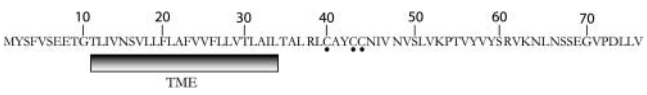

FIGURE 1 Complete sequence of SCoVE. The predicted TME used in the simulations is indicated (*shaded bar*). The corresponding transmembrane sequence used for other variants is shown in the alignment of Fig. 2. Three cysteines (*black circles*) C40, C43, and C44 are indicated, which are possible palmitoylation sites.

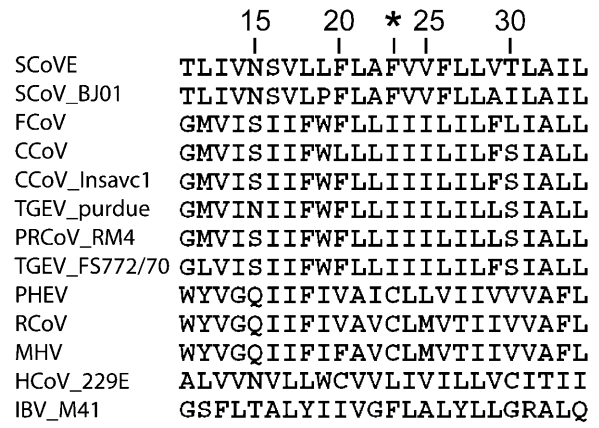

FIGURE 2 Sequences corresponding to the putative transmembrane segments of SARS coronavirus E protein and its homologous used in our molecular dynamics simulations. The column on the left indicates their abbreviated name. The complete name and corresponding Swiss-Prot entries are indicated in the Materials and Methods section. The numbering corresponds to SCoVE. The residue used to calculate the rotational orientation, ω_{23}, for the models in Figs. 3–5 is indicated by an asterisk.

The assignment of the transmembrane domain for each sequence was based on the hydrophilicity/surface probability plots (Kyte and Doolittle, 1982; Emini et al., 1985) and the transmembrane prediction on the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM) (Krogh et al., 2001). According to these predictors, the transmembrane region for the sequences spans ~24 residues and the following residues were used for the simulations: 11–34 for sequences SCoVE, SCoV_BJ01, RCoV, and MHV; 14–37 for sequences CCoV_Insavc1, THEV_purdue, PRCOV_RM4, TGEV_FS772/70, FCoV, and CCoV; 10–33 for sequence HcoV_229E; and 13–36 for sequences PHEV and IBV_M41, all containing the same number of residues, i.e., 24.

Global search molecular dynamics (GSMD) protocol

For the simulations we used a Hewlett-Packard Alpha SC45 Cluster containing 44 nodes. All calculations were performed using the Parallel Crystallography and NMR System (PCNS), the parallel-processing version of the Crystallography and NMR System (CNS Version 0.3) (Brunger et al., 1998), with united atom topology (Jorgensen and Tirado-Rives, 1988) explicitly describing only polar and aromatic hydrogen atoms. A global search was carried out in vacuo as described elsewhere (Adams et al., 1995), using CHI 1.1 (CNS Helical Interactions) and assuming a symmetrical interaction between the helices in the homooligomer.

Trials were carried out starting from either left or right crossing configurations. The helix tilt, β, was restrained to 0° and the helices were rotated a total of 350° about their long helical axes, in 10° increments. Henceforth, the simulation was repeated by increasing the helix tilt in discrete steps of 5°, up to 40°. We should point out, however, that this restraint is not completely strict, and the helix tilt at the end of the simulation can drift up to ±5° from the restrained value.

Three trials were carried out from each starting configuration using different initial random velocities. Increasing oligomeric sizes were examined, from 2 (dimers) to 6 (hexamers). Each protocol was repeated for up to 13 different sequences (Fig. 2). Hence, a total of 9 (tilt) × 36

(rotation) × 5 (size) × 13 (sequences) × 3 (repeats) × 2 (handedness) = 126,360 structures were produced and analyzed, i.e., 25,272 for each oligomeric size and 1,944 for each sequence and a given oligomeric size.

For each oligomeric size and helix tilt, clusters with a minimum number of structures (typically 10) were identified, where any structure belonging to a particular cluster was typically within 1.0 Å RMSD (root mean-square deviation) from any other structure within that cluster. The structures belonging to each cluster were averaged and subjected to energy minimization. These final structures were taken as the representative of the clusters and represented in the plots (see Figs. 3–5).

Analysis of the simulations

As previously (Briggs et al., 2001; Torres et al., 2002a), the results from the GSMD simulations were represented graphically by plotting each representative structure as a function of two parameters, helix tilt, β, and rotational orientation, ω of a specific residue, in the ordinate and abcissa axis, respectively. As described previously (Arkin et al., 1997), the rotational orientation angle ω is defined by the angle between a vector perpendicular to the helix axis, oriented toward the middle of the peptidic C=O bond of the residue, and a plane that contains both the helical axis and the normal to the bilayer. This angle is 0° when the residue is located in the direction of the tilt. For all representations in Figs. 3–5, the rotational orientation ω was defined relative to residue 23, indicated as ω_{23}, in the sequence SCovE (see *asterisk symbol* in Fig. 2), and its equivalent residue for other sequences. The tilt angle of the models, β, was taken as the average of the angles between each helix axis in the bundle and the bundle axis. The bundle axis, coincident with the normal to the bilayer, was calculated by CHI. The helix axis was calculated as a vector with starting and end points above and below a defined residue, where the points correspond to the geometric mean of the coordinates of the five α-carbons N-terminal and the five α-carbons C-terminal to the defined residue. Intersequence comparisons between low energy clusters were performed by calculating the RMSD between their α-carbon backbones. Fitting was performed using the program ProFit (http://www.bioinf.org.uk/software/profit). The energies calculated correspond to the total energy of the system, including both bonded, e.g., bond, angle, dihedral, improper, and nonbonded, i.e., van der Waals and electrostatic terms (Adams et al., 1995).

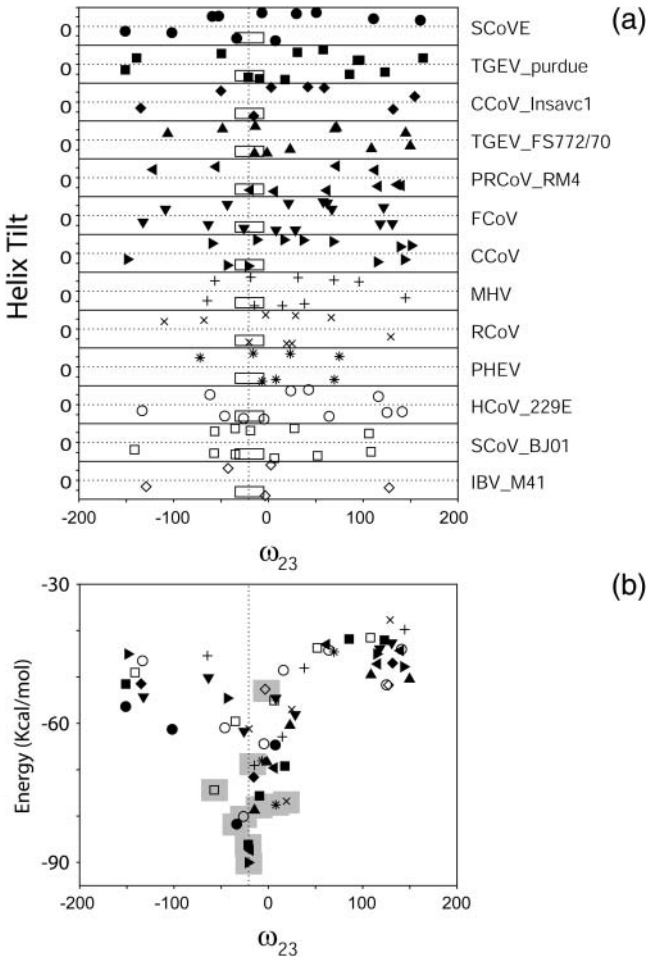

FIGURE 3 (*a*) Plot of helix tilt versus ω_{23} for the low energy models (each symbol represents one model) obtained after the GSMD simulations for a homodimeric model when restraining the helix tilt to 10°. For each sequence, the horizontal broken line separates left-handed (*symbols above the broken line*) from right-handed bundles (*symbols below the broken line*). The vertical broken line indicates the average orientation (at $\omega = -23°$) where the complete set was found (RMSD, 1.5 Å; $n = 10$ structures). The models inside the small rectangles are those forming a complete set. (*b*) The models in panel *a* are represented as a function of their energy (ordinate axis) and ω_{23}. The lowest energy models found in each sequence are indicated with a shaded rectangle.

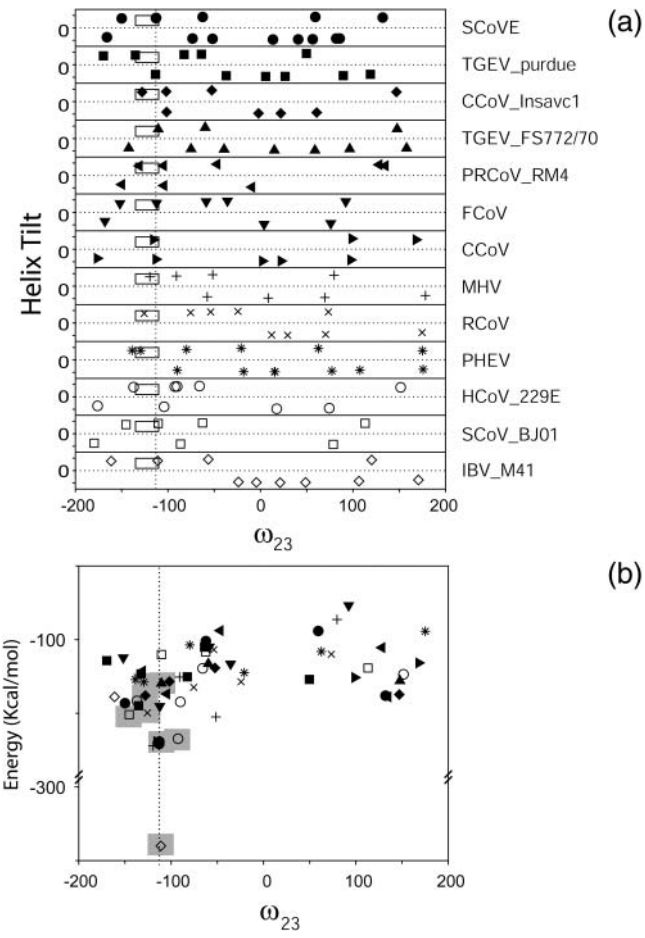

FIGURE 4 As in Fig. 3, but assuming a homotrimeric homooligomer. This figure only shows the results when the helix tilt was restrained to 35°. The vertical broken line indicates the orientation at $\omega = -113$, where the complete set was found (RMSD, 1 Å; $n = 10$).

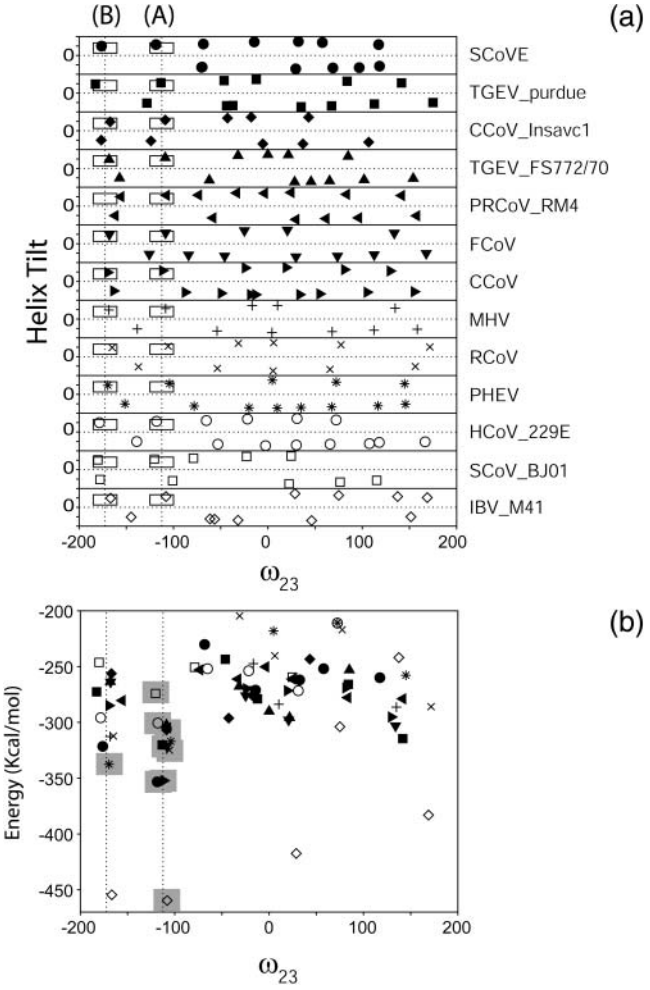

FIGURE 5 As in Fig. 3, but assuming a homopentameric homooligomer. Only restraining the helix tilt to 25° (*shown here*), a complete set was found (RMSD, 1Å; $n = 10$). The vertical broken lines indicate the orientation of the complete sets found, at $\omega = -121°$ (form *A*) and at $\omega = -176°$ (form *B*).

Synthesis of the transmembrane peptide of SARS protein E and SDS-PAGE electrophoresis

The peptide corresponding to the transmembrane helix of SARS protein E was synthesized in a Respep peptide synthesizer (Intavis Bioanalytical Instruments AG, Cologne, Germany), using standard solid-phase FMOC chemistry, from residue 9 to 35, and adding 2 lysines to both N- and C-ends, to improve solubility. The exact sequence used was KKTGTLIVNSVLLFLAFVVFLLVTLAILTKK, amidated and acylated at C- and N-termini, respectively. The peptide was cleaved from the resin with trifluoroacetic acid (TFA) and lyophilized. The lyophilized peptides were dissolved in trifluoroethanol (TFE), TFA and acetonitrile (1:1:4, v/v/v) (final peptide concentration ~5 mg/ml) and immediately injected to a 20-ml Juppiter 5 C4-300 column (Phenomenex, Cheshire, UK) equilibrated with H_2O. Peptide elution was achieved with a linear gradient to a final solvent composition of 10% H_2O, 90% acetonitrile, using a Waters 600 HPLC system. All solvents contained 0.1% (v/v) TFA. The resulting fractions were pooled and lyophilized. Peptide purity was confirmed by mass spectrometry.

The electrophoretic mobility of the peptide was assessed using SDS/PAGE. SDS sample buffer was added to the lyophilized peptide to a final concentration of 2 μg/μl. After vortexing for 1 min the sample was heated at 70°C for 5 min and loaded on a 15% SDS-PAGE gel (Tris Glycine). The loading volumes were 5, 10, and 20 μl. The sample was electrophoresed at room temperature at a constant voltage of 100 V for 30 min. After completion, the SDS/PAGE gel was first stained with Coomassie blue, followed by silver staining with the Silver stain-Plus kit (Bio-Rad).

RESULTS

Homodimer simulations

Fig. 3 (*panel a*) shows the results of the simulations when TME was assumed to be a homodimeric α-helical bundle. Only the results corresponding to a restrained helix tilt to 10° are shown, because no persistent models were found at any other helix tilt tested. The preserved configuration is right-handed (*models below the horizontal broken line*) and has an average orientation of $\beta = 12°$ and $\omega_{23} = -23°$ (*vertical broken line*). To guide the eye, the models consistent with this configuration for each of the sequences have been enclosed within a small rectangle. We note that helix tilt versus ω is just a convenient way of representation, and structures with up to 20° difference in the ω for a certain residue can in fact be very similar, e.g., for SCoVE ω_{23} is $-33°$, not $-23°$.

No structure within this set, which spans all sequences tested, was found to differ from any other in the same set by more than 1.5 Å Cα RMSD. This RMSD value is higher than that reported previously (below 1 Å RMSD) using the same method for various other homooligomers (Briggs et al., 2001), which casts some doubts on the relevance of this structure. However, one must take into account the low similarity between the transmembrane sequences used here (17%) compared to those used in previous work (more than 50%) (Briggs et al., 2001; Torres et al., 2002a). It is therefore possible that the high RMSD observed is due to the low similarity of the sequences used, which in turn may indicate that the structure represented by these sequences is not identical. In fact we have observed a smaller RMSD (1.15 Å) when using sequences from the same coronavirus group.

We can also assess the relevance of this model by observing the energy values obtained in each simulation (Torres et al., 2002a). If the model is correct, the lowest energy models for each sequence will tend to cluster around that particular conformation. Panel *b* in Fig. 3 shows that the lowest energy models (highlighted by *shading*) for each sequence cluster around $\omega = -23°$ (*vertical dotted line*), which is where the persistent conformation appears. We conclude therefore that protein E forms a homodimeric structure. Slices corresponding to this dimeric model for sequence SCoVE are represented in Fig. 6 (*left column*).

Homotrimer simulations

Fig. 4 shows the results of the simulations assuming a homotrimeric α-helical bundle when the helix tilt was

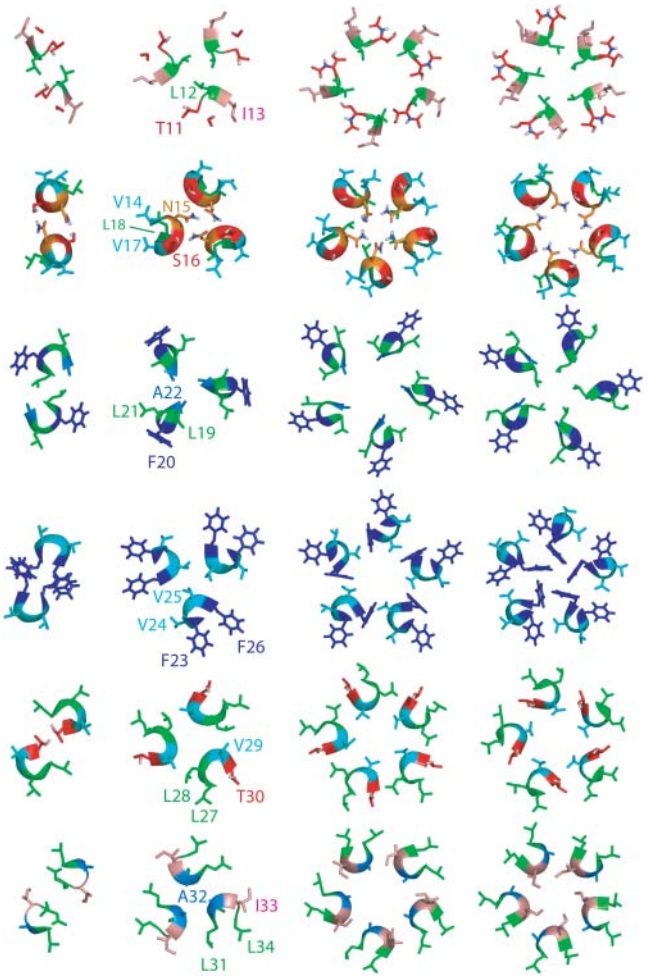

FIGURE 6 Columns from left to right: slices through the dimeric, trimeric, pentameric-form A and pentameric-form B of the transmembrane domain of SARS coronavirus E protein, i.e., sequence SCoVE in Fig. 2. Color code: L, green; V, cyan; I, salmon; A, marine; F, blue; N, orange; and S and T, red. For clarity's sake, the residue numbers are indicated only in one of the helices of the trimeric model. Note the central role of N15 for the three types of oligomers.

restrained to 35°. Only in this case a persistent left-handed conformation was found, at $\beta = 32°$, $\omega = -113°$. No structure within this complete set (*see symbols within small rectangles*) differed from any other in the same set by more than 1 Å Cα RMSD. As in the case of the dimer, the energy plot in panel *b* shows that the lowest energy model for each simulation-sequence appears at, or near, the ω representing the complete set. The slices corresponding to this trimeric model are presented in Fig. 6.

Homotetramer simulations

For the homotetramer, no complete set like those described for dimer and trimer could be found for any restrained helix tilt, even at 2 Å Cα RMSD. The results are not shown.

Homopentamer simulations

Fig. 5 (*panel a*) shows the results of the simulations assuming a homopentameric arrangement. In this case, only when the helix was restrained to 25° and in a left-handed configuration, not one, but two persistent models were found. One model (*A*) appeared at $\beta = 23°$, $\omega = -121°$ (*right vertical broken line*) and the other model (*B*) appeared at $\beta = 20°$, $\omega = -176°$ (*left vertical broken line*). No structure within each of these complete sets represented by models A and B differed from any other structure in the same set by more than 1.0 Å Cα RMSD. When we tried to determine which of the models was correct based on their energies (*panel b*) we found that, except for PHEV, all lowest energy models are close or near $\omega = -121$, i.e., form A, which is a strong indication that model A is the correct one and model B must be a false positive. Intriguingly however, for the outlier sequence PHEV, the lowest energy model is equivalent precisely to form B ($\omega = -176°$). We hypothesize that models A and B could represent closed (*low energy*) and open (*high energy*) forms of a channel (see Discussion). Slices through both structures are given in Fig. 6 (two columns on the *right*).

Homohexamer simulations

As for the tetramer, no complete set could be found for any restrained helix tilt, even at 2 Å Cα RMSD. The results are not shown. Higher order oligomers were not tested.

SDS-PAGE of the transmembrane domain of SARS protein E

To assess experimentally the aggregation state of coronavirus protein E, the synthetic transmembrane domain of SARS protein E (TME) was solubilized in SDS and electrophoresed (see Materials and Methods). At the three concentrations of peptide tested (Fig. 7, *lanes* 2–4), we could observe bands

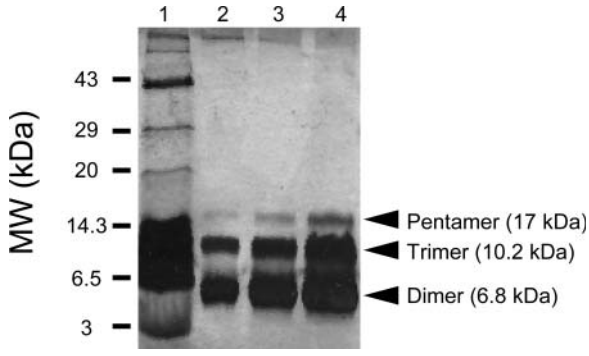

FIGURE 7 SDS-PAGE electrophoresis corresponding to the synthetic transmembrane peptide of SARS protein E. Lane 1 (*left*) shows the molecular weight markers. Lanes from 2 to 4: increasing load of peptide: 10, 20, and 40 μg, respectively. Arrows indicate the bands corresponding to the dimer, trimer, and pentameric forms of the peptide. The bands corresponding to the pentamer in lanes 2 and 3 were visible only after silver staining.

consistent with the presence of dimers, trimers and pentamers in SDS. Coomassie blue staining was sufficient for the most concentrated lane (*lane 4*), but after silver staining also lanes 2 and 3 showed the presence of the three oligomers. No other oligomeric form was detected.

DISCUSSION

After an exhaustive exploration of the conformational space of the transmembrane domain of protein E (TME), we have found that only a dimer ($\beta = 12°$, right-handed), a trimer ($\beta = 35°$, left-handed), and two pentamers ($\beta = 25°$, both left-handed) have been conserved by the conservative mutations appeared during evolution.

We note that, in contrast with previous work (Briggs et al., 2001; Kukol et al., 2002; Torres et al., 2002a), where we successfully obtained models in agreement with experimental data, no indication exists regarding the existence of a transmembrane α-helical homooligomer of coronavirus protein E in vivo or in vitro. Also, no structural data is available that permits to confirm or discard a given model. Could then these models have been conserved just by chance?

The first indication this is not the case is the extremely low probability that a model would survive all the conservative mutations present in 13 sequences with a similarity of only 17%. Clearly, the probability of finding a model only by chance decreases when the number of sequences analyzed is increased. Also, the lower the similarity between these sequences, the more stringent is the selection procedure, and the similarity between the sequences used here is as small as 17%, far below 50% used in previous work (Briggs et al., 2001).

The second indication that supports our prediction is based on the relative energies of the resulting models. If the correct model is the most stable, then it is expected that for every sequence the lowest energy model will be close to that represented by the complete set. Because of inaccuracies in the force-fields and other factors however, not all lowest energy models will have the same conformation, but they will cluster around the conformation of the correct structure, as confirmed previously in other proteins (Torres et al., 2002a). Consistent with this, the lowest energy models found in each aggregation state, dimeric, trimeric, or pentameric, for every sequence tested, cluster around the conformation of the persistent model.

The third indication is that, as we show in Fig. 7, the transmembrane domain of SARS protein E is sufficient to form dimeric, trimeric, and pentameric homooligomers in SDS micelles. Recent in vivo studies performed on the whole protein (Liao et al., 2004) also show dimers and trimers in nonreducing conditions, but only monomers in reducing conditions which prevent disulfide bond formation between monomers via the extramembrane cysteine residues (see Fig. 1). Overall, these combined results suggest that although the specificity in the interactions between monomers resides in the transmembrane domain, a dramatic effect on oligomerization is contributed by the extramembrane domain, specifically, but possibly not limited to, intermonomer disulfide bonds. In addition, the reported membrane permeabilizing activity of SCoVE (Liao et al., 2004) is consistent with our prediction of a pentameric bundle or pore, for which we find two close conformations, A and B. Although based on their energy values only model A appears to be correct, it is intriguing that the only persistent models found after a systematic search for a homopentamer should have the same handedness and helix tilt and be separated by a rotation of their helices of just 55°. We propose therefore that these two models could represent open and closed states of a channel, as both conformations should have been equally conserved during evolution. This is clearly reminiscent of phospholamban, where the possible existence of two transmembrane homopentameric models separated only by a rotation of their helices of ~40° has been discussed (Torres et al., 2001, 2002a).

The fourth and final indication is that we have predicted previously, using a method similar to the one described here, a transmembrane homotetrameric form for a component of the T-cell receptor, CD3-ζ (Torres et al., 2002b) for which only homodimers have been found experimentally. Recently, this prediction was partially confirmed by the observation of a homotetrameric form of the cytoplasmic domain of CD3-ζ (Sigalov et al., 2004) that could only be detected at very high concentrations. As both CD3-ζ and protein E are targeted to lipid rafts, it is possible that the local increase in concentration catalyzes the formation of many types of oligomers.

Another implication of this work is on protein E topology. Our results implicitly support a topology for protein E where N- and C-termini are in opposite sides of the membrane (Corse and Machamer, 2000). The latter authors showed unquestionably that the long hydrophilic tail (C-terminus) is facing the cytoplasm and therefore should be located in the inside of the virion envelope. Consistently, Raamsman et al. (2000) found that protein E was not digested after treating MHV particles with proteinase K in the absence of detergent. It was concluded that no part of protein E faces the virus exterior, although a question should be raised about the accessibility of such a small N-terminus to proteinase K (~10 amino acids, probably associated to the membrane). In fact, Yu et al. (1994) showed previously using an antibody against protein E in MHV that protein E was accessible from the surface of the virion envelope. Recently, Maeda et al. (2001) targeted a hydrophylic peptide (a flag) added to either the N- or the C-termini of protein E with antibodies and suggested that both N- and C-termini of the protein would reside in the cytoplasm, i.e., topologically equivalent to the virus lumen. These authors argued that an alternative explanation for the data reported by Corse and Machamer (2000) was that although both C and N are in the cytoplasmic domain, the N-terminus was simply too short to be targeted

by the antibody. Even more recently, in vitro biophysical studies (Arbely et al., 2004) have led to the suggestion that the putative TM domain of protein E forms a short hairpin that inserts only partially in the membrane.

We do not discard that more than one structural model for protein E can be present during the virion cycle, attending to its putative functional diversity. These different structural models could arise in different environments. For example, many viral envelope proteins contain putative palmitoylated sites and it is thought that this modification is important in protein-protein interactions during virus assembly. Palmitoylation may also induce raft partitioning, as shown for example in HIV-1 envelope glycoproteins (Bhattacharya et al., 2004). The cytoplasmic tail of SCovE, for example, contains one or more putative palmitoylation sites, and specifically, a ''double cysteine'' CC motif, which is a strong predictor of palmitoylation, as exemplified in various receptors (Bijlmakers and Marsh, 2003). In addition, in IBV (Corse and Machamer, 2002) and in MHV (Yu et al., 1994) protein E has already been shown to be palmitoylated. It is possible that this reversible covalent modification could trigger conformational changes critical for function. Physical interaction of this domain with the M protein is critical in the formation of the virion (Lim and Liu, 2001). An interesting possibility therefore is that palmitoylation and the subsequent close association of the cytoplasmic tail to the lipid bilayer may trigger a conformational change, or even the formation of a TM hairpin. Both in vitro and in vivo studies are currently being performed in our labs to test each of these hypotheses.

We are grateful to Paul D. Adams and Isaiah T. Arkin for kindly providing CHI, and John A. G. Briggs for helpful comments.

J.T. thanks the financial support of the Biomedical Research Council of Singapore and the facilities at the Bioinformatics Research Center of Nanyang Technological University.

REFERENCES

Adams, P. D., I. T. Arkin, D. M. Engelman, and A. T. Brunger. 1995. Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat. Struct. Biol.* 2:154–162.

An, S., C. J. Chen, X. Yu, J. L. Leibowitz, and S. Makino. 1999. Induction of apoptosis in murine coronavirus-infected cultured cells and demonstration of E protein as an apoptosis inducer. *J. Virol.* 73:7853–7859.

Arbely, E., Z. Khattari, G. Brotons, M. Akkawi, T. Salditt, and I. T. Arkin. 2004. A highly unusual palindromic transmembrane helical hairpin formed by SARS coronavirus E protein. *J. Mol. Biol.* 341:769–779.

Arkin, I. T., K. R. MacKenzie, and A. T. Brunger. 1997. Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers. *J. Am. Chem. Soc.* 119:8973–8980.

Baudoux, P., C. Carrat, L. Besnardeau, B. Charley, and H. Laude. 1998. Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *J. Virol.* 72:8636–8643.

Bhattacharya, J., P. J. Peters, and P. R. Clapham. 2004. Human immunodeficiency virus type 1 envelope glycoproteins that lack cytoplasmic domain cysteines: impact on association with membrane lipid rafts and incorporation onto budding virus particles. *J. Virol.* 78:5500–5506.

Bijlmakers, M. J., and M. Marsh. 2003. The on-off story of protein palmitoylation. *Trends Cell Biol.* 13:32–42.

Bos, E. C., W. Luytjes, H. V. van der Meulen, H. K. Koerten, and W. J. Spaan. 1996. The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus. *Virology.* 218:52–60.

Briggs, J. A., J. Torres, and I. T. Arkin. 2001. A new method to model membrane protein structure based on silent amino acid substitutions. *Proteins.* 44:370–375.

Brunger, A., P. D. Adams, G. M. Clore, W. L. P. Gross, R. W. Grosse-Kunstleve, J. S. Jiung, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1998. Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D.* 54:905–921.

Chen, C. J., S. An, and S. Makino. 2001. Induction of apoptosis in murine coronavirus-infected 17Cl-1 cells. *Adv. Exp. Med. Biol.* 494:615–620.

Corse, E., and C. E. Machamer. 2000. Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. *J. Virol.* 74:4319–4326.

Corse, E., and C. E. Machamer. 2002. The cytoplasmic tail of infectious bronchitis virus E protein directs Golgi targeting. *J. Virol.* 76:1273–1284.

Emini, E. A., J. V. Hughes, D. S. Perlow, and J. Boger. 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J. Virol.* 55:836–839.

Fischer, F., C. F. Stegen, P. S. Masters, and W. A. Samsonoff. 1998. Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly. *J. Virol.* 72:7885–7894.

Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS (optimized potentials for liquid simulations) potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

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

Kukol, A., J. Torres, and I. T. Arkin. 2002. A structure for the trimeric MHC class II-associated invariant chain transmembrane domain. *J. Mol. Biol.* 320:1109–1117.

Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–132.

Liao, Y., J. Lescar, J. P. Tam, and D. X. Liu. 2004. Expression of SARS-coronavirus envelope protein in *Escherichia coli* cells alters membrane permeability. *Biochem. Biophys. Res. Commun.* 325:374–380.

Lim, K. P., and D. X. Liu. 2001. The missing link in coronavirus assembly. Retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-Golgi compartments and physical interaction between the envelope and membrane proteins. *J. Biol. Chem.* 276:17515–17523.

Maeda, J., J. F. Repass, A. Maeda, and S. Makino. 2001. Membrane topology of coronavirus E protein. *Virology.* 281:163–169.

Narayanan, K., A. Maeda, J. Maeda, and S. Makino. 2000. Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. *J. Virol.* 74:8127–8134.

Opstelten, D. J., M. J. Raamsman, K. Wolfs, M. C. Horzinek, and P. J. Rottier. 1995. Envelope glycoprotein interactions in coronavirus assembly. *J. Cell Biol.* 131:339–349.

Raamsman, M. J., J. K. Locker, A. de Hooge, A. A. de Vries, G. Griffiths, H. Vennema, and P. J. Rottier. 2000. Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E. *J. Virol.* 74:2333–2342.

Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez,

S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*. 300:1394–1399.

Sigalov, A., D. Aivazian, and L. Stern. 2004. Homooligomerization of the cytoplasmic domain of the T cell receptor zeta chain and of other proteins containing the immunoreceptor tyrosine-based activation motif. *Biochemistry*. 43:2049–2061.

Torres, J., J. A. Briggs, and I. T. Arkin. 2002a. Contribution of energy values to the analysis of global searching molecular dynamics simulations of transmembrane helical bundles. *Biophys. J*. 82:3063–3071.

Torres, J., J. A. Briggs, and I. T. Arkin. 2002b. Convergence of experimental, computational and evolutionary approaches predicts the presence of a tetrameric form for CD3-zeta. *J. Mol. Biol*. 316:375–384.

Torres, J., A. Kukol, and I. T. Arkin. 2001. Mapping the energy surface of transmembrane helix-helix interactions. *Biophys. J*. 81:2681–2692.

Vennema, H., G. J. Godeke, J. W. Rossen, W. F. Voorhout, M. C. Horzinek, D. J. Opstelten, and P. J. Rottier. 1996. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J*. 15:2020–2028.

Yu, X., W. Bi, S. R. Weiss, and J. L. Leibowitz. 1994. Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Virology*. 202: 1018–1023.