Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Protein Expression and Purification 85 (2012) 133–141

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression and purification of coronavirus envelope proteins using a modified β-barrel construct

Krupakar Parthasarathy, Huang Lu, Wahyu Surya, Ardcharaporn Vararattanavech, Konstantin Pervushin, Jaume Torres *

School of Biological Sciences, Nanyang Technological University, Singapore, Singapore

ARTICLE INFO

Article history:
Received 13 June 2012
and in revised form 11 July 2012
Available online 20 July 2012

Keywords:
Membrane protein
β-Barrel fusion protein
Coronavirus
Envelope proteins
Cysteines
Cyanogen bromide

ABSTRACT

Coronavirus envelope (E) proteins are short (~100 residues) polypeptides that contain at least one transmembrane (TM) domain and a cluster of 2–3 juxtamembrane cysteines. These proteins are involved in viral morphogenesis and tropism, and their absence leads in some cases to aberrant virions, or to viral attenuation. In common to other viroporins, coronavirus envelope proteins increase membrane permeability to ions. Although an NMR-based model for the TM domain of the E protein in the severe acute respiratory syndrome virus (SARS-CoV E) has been reported, structural data and biophysical studies of full length E proteins are not available because efficient expression and purification methods for these proteins are lacking. Herein we have used a novel fusion protein consisting of a modified β-barrel to purify both wild type and cysteine-less mutants of two representatives of coronavirus E proteins: the shortest (76 residues), from SARS-CoV E, and one of the longest (109 residues), from the infectious bronchitis virus (IBV E). The fusion construct was subsequently cleaved with cyanogen bromide and all polypeptides were obtained with high purity. This is an approach that can be used in other difficult hydrophobic peptides.

Introduction

Envelope (E) proteins are found in coronaviruses (family *Coronaviridae*, genus *Coronavirus* [1]), a group of enveloped viruses that cause common cold in humans and a variety of lethal diseases in birds and mammals [2]. The species in the genus *Coronavirus* has been organized into three groups [3]. Group 1 includes the porcine *Transmissible gastroenteritis virus* (TGEV[1]) and *Human coronaviruses 229E* (HCoV-229E) or NL63 (HCoV-NL63), Group 2 includes *Murine hepatitis virus* (MHV) and *Human coronavirus OC43* (HCoV–OC43). Group 3 includes the avian *Infectious bronchitis virus* (IBV) and the Turkey coronavirus (TCoV). The virus responsible for the severe acute respiratory syndrome (SARS-CoV), a respiratory disease in humans [4], is close to group 2 [5].

Coronavirus E proteins are minor components in virions, with varying lenghts of around ~100 amino acids. All of them have at least one predicted α-helical transmembrane (TM) domain followed by a juxtamembrane cluster of 2–3 cysteines. The TM domain is flanked by a short hydrophilic N-terminus (~10 residues) and a C-terminal tail (39–76 residues).

E proteins are not essential for *in vitro* or *in vivo* coronavirus replication, but their absence results in attenuated viruses, as shown for SARS-CoV [6] and others [7–12]. In some coronaviruses, absence of E protein results in deletereous morphological changes in the virion [13–16]. E proteins has also been shown to have membrane permeabilizing [17] and ion channel [18–20] activity, although the purpose of this is not known.

Biophysical or structural studies so far have been performed using synthetic peptides corresponding to the predicted α-helical TM domain [17,21,22] or with synthetic full length protein [23,24]. In the latter case, however, the peptides were only partially purified and contained truncations or deletions. Thus, despite the evident biological interest of coronavirus E proteins, to our knowledge, none of the envelope coronavirus proteins has been efficiently expressed and purified.

Our initial attempts to overexpress SARS-CoV E, using poly-Histidine, GST, ketosteroid isomerase (KSI) [25,26] or maltose binding

* Corresponding author. Address: School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore. Fax: +65 6791 3856.
E-mail address: jtorres@ntu.edu.sg (J. Torres).

[1] *Abbreviations used:* BBP, beta barrel protein; $BBP_{\Delta M}$, BBP where internal methionines has been mutated to Alanine; CNBr, cyanogen bromide; CoV, coronavirus; DMPC, 1,2-dimyristoylphosphocholine; DPhPC, diphytanoylphosphatidylcholine; E, envelope; IBV, infectious bronchitis virus; MALDI-TOF/MS, matrix-assisted laser desorption/ionization-time of flight mass spectrometry; MHV, mouse hepatitis virus; RP-HPLC, reverse phase high performance liquid chromatography; SARS, severe acute respiratory syndrome; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TE, Tris/EDTA; TGEV, transmissible gastroenteritis virus; TFA, trifluoroacetic acid; OmpA, outer membrane protein A; HABA, 2-(4-hydroxyphenylazo) benzoic acid.

1046-5928/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2012.07.005

protein (MBP) fusion tags were not successful. While the first three resulted in very low expression levels, MBP could not be efficiently purified from SARS-CoV E (results not shown).

Thus, we used a novel construct in which the E protein is N-terminally fused to an integral membrane protein, a modified eight strand β-barrel protein derived from the *Escherichia coli* outer membrane protein A (OmpA) [27]. This β-barrel, which has shortened extramembrane loops, has been previously characterized by NMR [28], and is referred herein as β-barrel protein platform (BBP). A 6His-tag was added to the N-terminus of BBP to further facilitate purification of the fusion construct. BBP was linked to the N-terminus of the E protein with a flexible linker region which contained the restriction site for factor Xa.

Although we obtained high expression levels (20 mg/L), proteolysis with factor Xa was not effective. Thus, the construct 6His-BBP-E was cleaved chemically with cyanogen bromide (CNBr), which cleaves polypeptides at methionine residues [29–31]. CNBr therefore cannot be used when internal methionines are present in the target protein, which fortunately is not the case for neither SARS-CoV E or IBV E. However, the β-barrel used as fusion protein contains four internal Met residues, and cleavage of the fusion construct with CNBr would produce in this case many fragments that could complicate purification. To avoid this, all four internal BBP Met residues were mutated to Ala (mutant $BBP_{\Delta Met}$).

To demonstrate that this method is efficient, we have expressed and purified two representatives of coronavirus E proteins, the shortest, SARS-CoV E (76 aa, 8.3 kDa), and the longest, IBV E (108 aa, 12.3 kDa).

These two proteins, like other coronavirus envelope proteins, contain several (2–3) juxtamembrane cysteines of as yet unknown function that are conserved. It is known, however, that they can be palmitoylated *in vivo* [21,32,33], and that removal of these cysteines in MHV E resulted in deformed viruses [34,35]. The study of the role of these cysteines is therefore an objective in their own right. In addition, these cysteines could potentially produce inter or intra-molecular disulfide bonds, leading to sample heterogeneity. Thus, the potential negative effect of disulfide bond formation in the final samples was examined by also purifying these proteins without cysteines. The latter samples may also be more amenable for structural studies.

Both envelope proteins and their cysteine-less mutants could be purified by reverse phase high performance liquid chromatography (RP-HPLC). This method should facilitate future detailed biophysical and structural studies of coronavirus envelope proteins. An additional benefit of this fusion protein is that BBP was originally engineered by inserting additional 12 amino acids in a loop region of the barrel, corresponding to loop 3 of calmodulin from *X. laevis*. This modified loop binds lanthanides [36,37], and the use of BBP can enable measurement of residual dipolar couplings (RDC) in membrane proteins [28].

Materials and methods

Constructs

pET3a/$^{6H}BBP_{\Delta M}$ expression vector was modified from the plasmid pET3a-BBP described in a previous work [28]. The mutated BBP ($BBP_{\Delta M}$) was generated with its four internal methionine residues replaced with alanine using the QuickChange® SDM kit (Stratagene) with the relevant primers. Additionally, a hexa-His tag was introduced at the N-terminus of BBP to facilitate purification. The forward and reverse primers used to amplify the target gene SARS-CoV E and IBV-E are as follows:

SARS-CoV E (NCBI Reference Sequence: NC_004718.3, protein ID NP_828854.1): forward primer 5′CAGGATCCATCGAAGGTCGTATGTACTCATTCG3′ and reverse primer 5′GAGGATCCTCAGACCAGAAGATCAGGAAC3′.

IBV E (NCBI Reference Sequence: NC_004718.3, protein ID AAU14251): forward primer 5′ CAGGATCCATCGAAGGTCGTATGTTTAATTTATTCC3′ and reverse primer 5′ GAGGATCCTTAGATATCATCCACCTCTAATAG 3′.

Both forward and reverse primers contain *BamH*I restriction sites. Factor Xa cleavage sequence was introduced at the N-terminus of the genes. The amplified inserts were digested with *BamH*I and then ligated to the similarly digested pET3a/$^{6H}BBP_{\Delta M}$ expression vector. The identity of the constructs was confirmed by DNA sequencing.

Mutants of SARS-CoV E and IBV-E constructs at cysteine residues were generated using the QuickChange® SDM kit (Stratagene). In SARS-CoV E, three cysteine residues at position 40, 43 and 44 (represented by the first, second and third letters, respectively) were mutated to serine or alanine, generating triple Cys mutants (SSS and AAA). In IBV E, the two cysteine residues at positions 44 and 45 were mutated to Alanine, generating a double-cysteine mutant (AA).

Overexpression and purification of the constructs containing SARS-CoV E and IBV E

The plasmid containing the desired construct was transformed into the *E. coli* strain BL21- CodonPlus (DE3)-RIL (Stratagene). One colony was inoculated into LB media (3 ml) containing 50 μg/ml of ampicillin and 34 μg/ml of chloramphenicol and incubated overnight at 37 °C with shaking at 200–225 rpm. The culture was inoculated into 500 ml LB media in a 2.5 L flask containing the same concentrations of antibiotics. Protein expression was induced with 1 mM isopropyl β-*d*-thiogalactopyranoside (IPTG) at an OD_{600} of 0.5–0.6. Expression of IBV_{AA} mutant was much lower than the wild type, and best expression was obtained by using 0.4 mM IPTG.

Cells were harvested by centrifugation when cell growth reached its stationary phase. Expression was checked by running the whole-cell lysate on SDS–PAGE. The fusion protein was found in inclusion bodies. For a 0.5 L culture, the cell pellet obtained after centrifugation was first washed with 20 mM Tris–HCl and 5 mM EDTA (TE buffer, pH 8) and then resuspended in 20 ml of 20 mM Tris–HCl and 5 mM EDTA. The resuspended cells was lysed by sonicating for 10–15 min in ice, followed by microfluidization 5–10 times and centrifugation at 8,000 rpm at 4 °C for 20 min. The white pellet was collected and resuspended in 20 ml TE buffer (pH 8) containing 2% Triton X-100, shaken at 37 °C for 1 h, and centrifuged at 8000 rpm at 4 °C for 30 min. The pellet was resuspended in TE buffer, incubated at 37 °C for 1 h and centrifuged at 8000 rpm at 4 °C for 20 min.

The pellet was then resuspended completely in the binding buffer (8 M urea, 5 mM imidazole in 20 mM Tris pH 8.0) and centrifuged at 12,000 rpm at 4 °C for 30 min. The protein was allowed to bind overnight onto Ni-NTA resin (Bio-rad) pre-equilibrated with the binding buffer. After incubating, the resin was washed with 5 column volumes of washing buffer (8 M urea, 20 mM imidazole in 20 mM Tris pH 8.0) followed by elution (elution buffer: 8 M urea, 250 mM imidazole in 20 mM Tris pH 8.0). All fractions were analyzed by SDS–PAGE. The desired fraction was dialysed to remove urea, and the precipitate was collected and lyophilized for CNBr cleavage.

The same protocol was followed for the expression of the ^{15}N labeled SARS-CoV E, although with some modifications. Initially, the cells were grown in normal LB medium and after the OD_{600} reached 0.6, the cells were harvested by centrifugation and resuspended in minimal media. After two cycles of resuspension and centrifugation, cells were inoculated in M9 minimal medium enriched with ^{15}N labeled ammonium chloride (1 g/L, >99% ^{15}N)

(Cambridge Isotopes, USA). In addition, the medium was supplied with 1% basic mineral elements solution (Sigma chemicals) which includes trace elements and vitamins.

MALDI-TOF/MS (Applied Biosystems) was used for identification of the cleavage products. The protein was solubilized in 60% isopropanol, 35% acetonitrile, 5% water and 0.1% TFA and deposited onto a matrix of 2-(4-hydroxyphenylazo) benzoic acid (HABA).

CNBr cleavage of the fusion protein

The lyophilized powder of purified construct was solubilized in 80% formic acid (to a few ml). In a different flask, five-fold molar excess of CNBr versus methionine content of the construct was also dissolved in 80% formic acid. After mixing both solutions the reaction proceeded in the dark under anaerobic conditions [38]. After 12–15 h, the solution was diluted and dialysed against water. The precipitate was lyophilized and solubilized by addition of 10 μL of trifluoroacetic acid (TFA). Acetonitrile was added to reach a 99.9:0.1 (acetonitrile to TFA) volume ratio. The solubilized CNBr cleavage products were injected into a Jupiter 5 μ C4 300 Å analytical column with a linear gradient of water (Solvent A: 95% H_2O, 5% acetonitrile, 0.1% TFA), and 2-propanol (Solvent B: 80% 2-propanol, 20% acetonitrile, 0.1% TFA). The presence and purity of SARS-CoV E or IBV E in the HPLC peaks was assessed by MALDI-TOF/MS using a HABA matrix in a 1:1 (v/v) ratio with HPLC fractions. Fractions with peaks matching the molecular weight of the protein were concentrated using a Scanvac centrifuge (Chemoscience) to remove excess organic solvents, diluted with H_2O and lyophilized. Lyophilized SARS-CoV E or IBV E powder was solubilized in 20–30 μl of 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP), and the concentration of the protein was estimated using its absorbance at 280 nm.

Electrophoresis of SARS-CoV E and IBV E proteins

SDS solubilizing buffer was added to the lyophilized protein to a final concentration of 1 mg/mL, with or without DTT. After mixing for 5 min, the sample was heated at 65 °C for 5 min and loaded. The loading volume was 10 μl (10 μg of peptide). Four percent to 12% NuPAGE Bis-tris gel (Invitrogen) was run at 150 V for 1 h with 2-(*N*-morpholino)-ethanesulfonic acid (MES)-SDS running buffer. After completion, the SDS–PAGE gel was stained with Coomassie blue.

For PFO electrophoresis, peptides were dried from HFIP and dissolved in sample buffer containing 4% PFO to ~1 mg/ml. Samples were heated at 65 °C for 4 min and loaded on a 4–12% NuPAGE®-Bis-tris gel (Invitrogen). The gel was electrophoresed at 80 V for 2–3 h with 2-(*N*-morpholino)-ethanesulfonic acid (MES) PFO running buffer. After completion, the gel was washed twice with water, followed by Coomassie blue staining.

In-gel tryptic digestion

Protein identity was confirmed by tryptic digestion after running SDS–PAGE. The protein band corresponding to SARS-CoV E or IBV E protein was excised and subjected to in-gel trypsin digestion. The gel fragments were diced into small fragments (1 mm^2) and washed three times with 200 μl 50% acetonitrile and 25 mM ammonium bicarbonate to remove the Coomassie blue dye. Gel fragments were then soaked in 100% acetonitrile for 10 min. Using the gel loading tip, acetonitrile was removed and residual acetonitrile was removed by Speedvac. Prior to enzymatic digestion, gel fragments were reduced with 10 mM DTT in 25 mM ammonium bicarbonate at 60 °C for 30 min. Alkylation was then performed with 55 mM iodoacetamide in 25 mM ammonium bicarbonate for 30 min at room temperature in the dark. The reduced and alkylated gel slices were rehydrated in 20 μl of 50 mM ammonium bicarbonate, pH 8.3, containing 200 ng of trypsin. Once this solution was fully absorbed by the gel slices, enzyme-free ammonium bicarbonate buffer was added until the gel pieces were completely covered. The samples were digested for 18 h at 37 °C [39]. After in gel tryptic digestion, the peptides were extracted and analyzed using MALDI-TOF/TOF-MS as described previously [40].

Circular dichroism (CD) and infrared spectroscopy

CD spectra were measured on a Chirascan CD Spectrometer (Applied Photophysics, UK). Data were collected at 20 °C at 1 nm interval, from 190 nm to 240 nm, on a 1 mm pathlength cell. SARS-CoV E and detergents were first dissolved in HFIP and dried, followed by addition of 1 mM phosphate buffer at pH 7. Final detergent concentrations were: 13 mM SDS and 10 mM DPC. The final protein concentration was approximately 10 μM. Spectra were smoothed with a Salvitsky-Golay algorithm.

Infrared spectra were acquired as described previously on a Nicolet Nexus spectrometer (Madison, USA) [20]. The protein was incorporated in multilamellar liposomes by first dissolving a dry mixture of lipid and lyophilized peptide (50:1 M ratio) in HFIP. The solution was dried by a stream of N_2, and the resulting peptide-lipid film was dissolved in phosphate buffer (1 mM, pH 7) followed by sonication. The compositions of lipid mixtures used were as follows: DMPC alone, DMPC:DMPG (1:1 M ratio) and POPC:SM:CH (2:1:1, w/w/w). The liposome solution was deposited on a germanium internal reflection element (IRE) and dried slowly by evaporation.

NMR of ^{15}N labeled SARS-CoV E protein

NMR experiments were performed using a Bruker Avance-II 600 or 700 NMR spectrometers, equipped with cryogenic probes (Bruker BioSpin, GmbH), with three channels and triple resonance

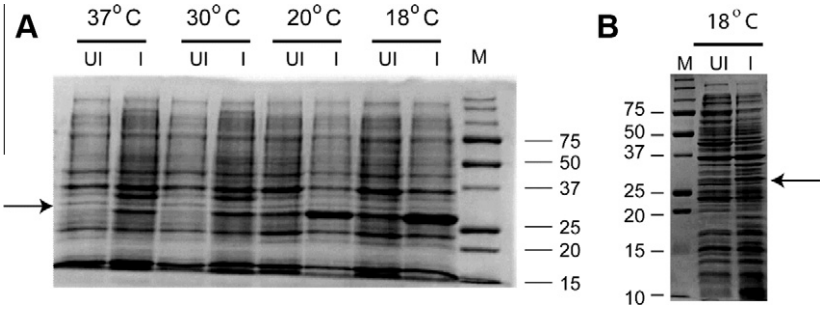

Fig. 1. Coomassie-stained SDS–PAGE (12%) showing the expression profile of SARS-CoV E (A) and IBV E (B) when fused to $BBP_{\Delta M}$ (6His-$BBP_{\Delta M}$-E construct). In (A), several temperatures of expression tested are indicated; M, molecular mass standards; UI, uninduced; I, induced. The arrow indicates the band corresponding to the fusion construct. In (B), only expression at 18 °C is shown.

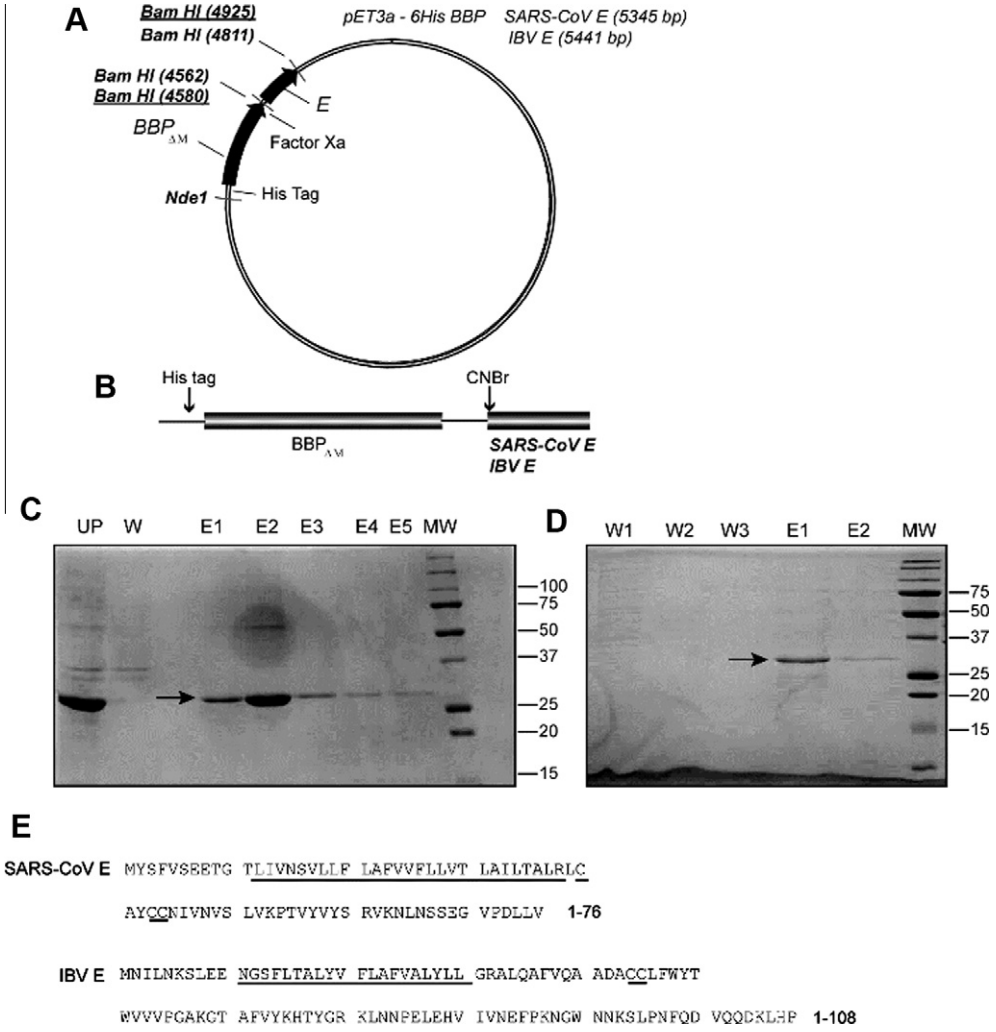

Fig. 2. Purification of fusion construct containing SARS-CoV E and IBV E. (A) schematic representation of the plasmid containing the construct, with restriction sites, histidine tag, mutated BBP without internal methionines ($BBP_{\Delta M}$) and the target SARS-CoV E or IBV E; (B) simplified representation of the fusion construct with CNBr cleavage site; (C) Coomassie-stained SDS–PAGE gel (12%) with Ni-NTA affinity chromatography purification profile of the fusion construct containing SARS-CoV E: UP, unpurified sample; W, washing; E, elutions; MW, molecular weight markers. The arrow indicates the band corresponding to the fusion protein; (D) same as (C), for the fusion construct containing IBV E: W1, flow-through after overnight incubation; W2 and W3, flow-through after washing buffer; E1, first elution; E2, second elution. Arrow shows fusion protein (30.3 kDa); (E) sequences of SARS-CoV E (accession number NP_828854) and IBV E (accession number AAU14251). The predicted α-helical TM domain is underlined.

probes with shielded *z*-gradient coils. The carrier position was set to 4.70 ppm for ^{1}H and 119 ppm for ^{15}N. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as internal reference for ^{1}H nuclei. The chemical shifts of ^{15}N nuclei were indirectly calculated from the ^{1}H chemical shifts. SARS-CoV E protein was screened for solubilization in detergents DPC and SDS in phosphate buffer at pH 5.5 at various protein to detergent molar ratios, from 1:100 to 1:150. DTT was also used to reduce possible disulphide bonds.

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed at 25 °C using a Beckman XL-I analytical ultracentrifuge [41] in the presence or absence of 6 mM of reducing agent tris(2-carboxyethyl)phosphine (TCEP). Absorbance was measured at 230 nm, and for SARS-CoV E we used the estimated extinction coefficient ε_{230} = 27,000 M^{-1} cm^{-1}. The buffer composition was 50 mM Tris, 100 mM NaCl, pH 7.3 and 5 mM C14 betaine (c.m.c., 0.1 mM). To match the density of the detergent, D_2O was added to the buffer to a final volume ratio of 29.4% [42]. The density-matched buffer was added to the dry peptide. The molar ratio peptide:detergent was 1:25, 1:50 and 1:100 for inner, medium and outer cell, respectively. The samples were centrifuged in three-compartment carbon-epoxy centerpieces with quartz windows for lengths of time sufficient to achieve equilibrium, tested with WinMatch. This typically was achieved ~26 h after reaching the speeds at 38,000, 45,000 or 48,000 rpm. Data was processed and analyzed using Sedfit and Sedphat [43]. The monomeric molecular mass of the protein and its partial specific volume were calculated with SEDNTERP [44] after correction for partial hydrogen/deuterium exchange [41]: 8,229.84 Da and 0.7687 cm^3/g. The log plot (ln A vs. $r^2 - r_0^2$) was not linear, indicating the presence of more than one species at these peptide-to-detergent ratios. The data was fitted to several monomer-Nmer models.

For IBV E, the conditions were identical to that of SARS-CoV E except that the concentration of detergent (C14 betaine) was 15 mM. The protein:detergent molar ratio was 1:300, 1:450 and 1:600. The samples were centrifuged at 28,000, 34,500, 42,000 or 48,000 rpm.

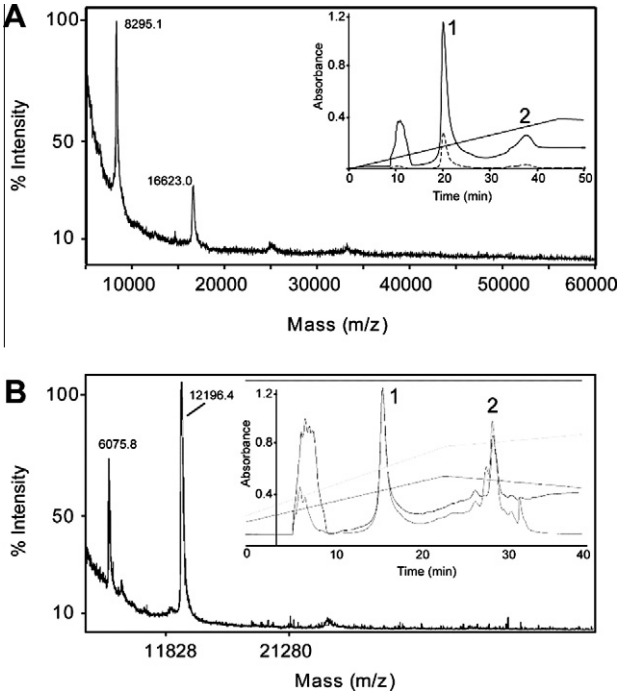

Fig. 3. MALDI-TOF/MS mass spectra for SARS-CoV E and IBV E. (A) mass spectrum of SARS-CoV E protein. The calculated and observed molecular weights are 8.230 and 8,295.1. The peak at 16.623 corresponds to the SARS-CoV E dimer. The insert shows the elution profile monitored at 220 nm (solid line) and 272 nm (broken line) for $BBP_{\Delta M}$ (peak 1) and SARS-CoV E (peak 2) after CNBr cleavage; (B) Same for IBV E (peak at 12,196.4 Da), and its double-protonated form (6,075.8 Da). This sample corresponds to peak 2 in the HPLC chromatogram (see insert).

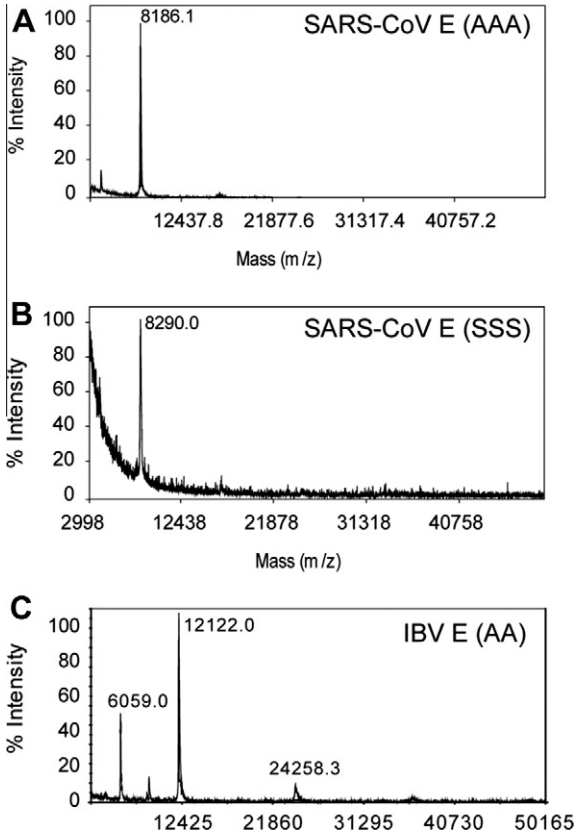

Fig. 4. MALDI-TOF/MS spectra of the purified mutants of SARS-CoV E and IBV E. The calculated mass of the monomer is 8,133 Da for the SARS-CoV E AAA mutant (A), 8,181 Da for the SSS mutant (B), and 12,116 Da for IBV E AA mutant (C). The band at ~6 kDa in (C) is the doubly protonated species, and the band at ~24 kDa is a dimeric form. The three letters represent the three cysteine positions C40, C43, C44 in SARS-CoV E, occupied by S (serine) or A (alanine). In IBV E, the two cysteines are substituted to Ala.

Channel activity

For the preparation of SARS-CoV E proteoliposomes, the protein was reconstituted into liposomes formed by DPhPC and cholesterol (10:1 M ratio, Avanti Polar Lipids), at a 1:100 protein/lipid (mol/mol) ratio. Giant unilamellar vesicles were prepared by electroformation on a Nanion Vesicle Prep Pro setup (Nanion Technologies GmbH, Munich, Germany). Electrophysiological recordings were carried out with a Port-a-Patch automated patch clamp system (Nanion Technologies GmbH, Munich, Germany). First, a planar bilayer was obtained by applying 5 µl of GUV solution onto the borosilicate glass chips, ~1 µm aperture, in presence of a slightly negative pressure (−20 mbar). The formation of the planar lipid bilayer on the aperture provided a seal resistance of >10 GΩ. Then, proteoliposomes containing the protein were pipetted onto the planar lipid bilayer. Currents were amplified with Axon Patch 200B amplifier (Axon instruments, USA). The data were acquired by using Axon Digidata 1322A and analyzed with Clampfit (Axon instruments, USA). The experiment was performed in a symmetric solution of 150 mM NaCl, 5 mM HEPES at pH 5.1, varying the voltages across the lipid bilayer.

Results and discussion

Expression and purification of SARS-CoV E

To determine the optimal temperature for expression, yield was tested at 37, 30, 20 and 18 °C. For both SARS-CoV E (Fig. 1A) and IBV E (Fig. 1B), highest expression was observed at 18 °C. Elution of the fusion construct containing SARS-CoV E (Fig. 2A and B), from a Ni-NTA column with 250 mM imidazole produced a relatively pure sample (Fig. 2C, lanes E1 and E2). The yield of fusion construct after this step was approximately 20 mg/L of culture. After CNBr cleavage, the HPLC elution profile showed two distinct peaks labeled 1 and 2, respectively (Fig. 3A, insert), eluting at 38% and 72% of solvent B. MALDI-TOF/MS analysis of peak 1 produced a mass of 18,415.1 Da corresponding to $BBP_{\Delta M}$ (not shown), and peak 2 produced a main peak at 8,295.1 Da (Fig. 3A), which corresponds to the monomeric SARS-CoV E (calculated mass 8,230 Da). Additional peaks at multiples of the monomeric form were observed, e.g., 16,623 Da, likely to represent dimers of SARS-CoV E.

The ^{15}N labeled SARS-CoV E protein was observed at 8,436 Da (see Supplementary File, Fig. S1B), whereas the calculated molecular weight was 8,316 Da. Like in unlabeled protein, multiples of this peak at 17,016, 25,532 and 34,901 Da were observed, corresponding to higher oligomers. We note that in both labeled and unlabeled protein, not only the observed molecular weight was higher than the calculated one, but the value obtained fluctuated between measurements. For example, in a total of five measurements, the average was 8,339 ± 23 Da for unlabeled SARS-CoV E protein. These alliquotes originated from the same volume, and the machine was properly calibrated. In any case, we have further confirmed the identity of the purified protein by in-gel tryptic digestion and identified the fragments by MALDI-TOF (see Supplementary File, Fig. S2).

The final yield of non-labeled SARS-CoV E protein was 4 mg/L of culture, which is comparable or higher than other methods of expression for this class of proteins [45], whereas the yield for ^{15}N-labeled protein was ~15 mg/L. For cysteine-less mutants of SARS-CoV E, expression and purification results were similar (not

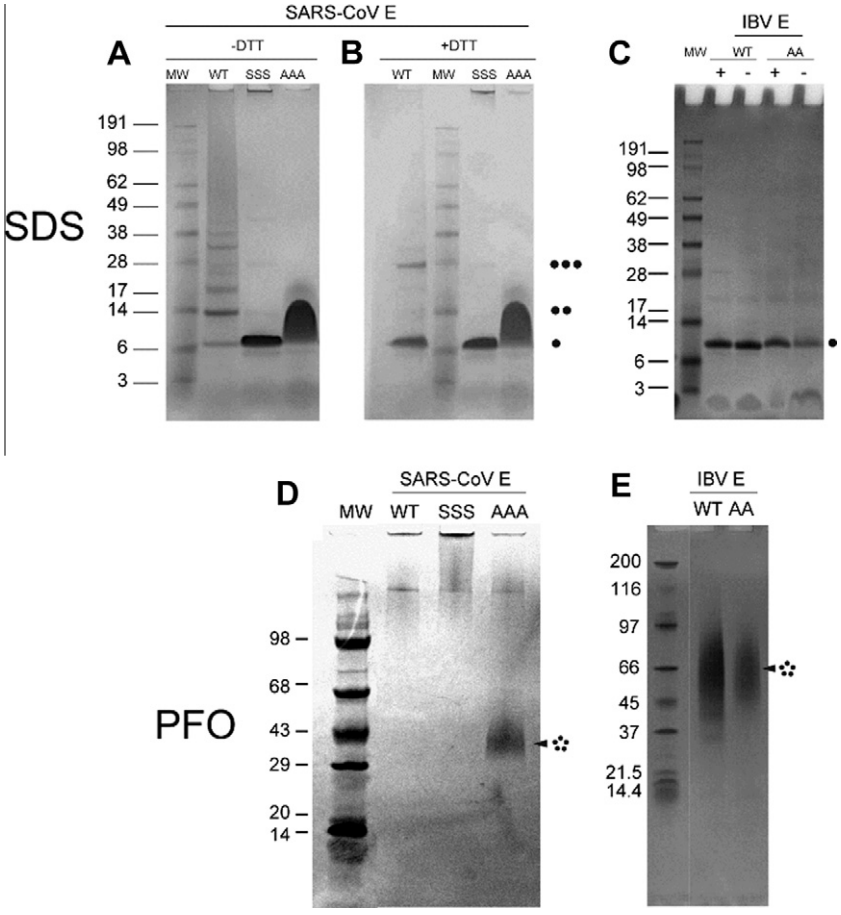

Fig. 5. SDS and PFO–PAGE electrophoresis of SARS–CoV E and IBV E. (A) SARS-CoV E wild type (WT) and triple-cysteine mutants (AAA, SSS) in the absence of DTT in gel 4–12% Nu–PAGE in MES/SDS buffer; (B) same as (A), in presence of DTT. The lane containing the molecular weight markers (MW) is indicated. The oligomeric size is indicated by black circles (●); (C) IBV E (WT) and double cysteine (AA) mutant with (+) or without (-) DTT; (D) SARS-CoV E and mutants in 4–12% PFO Nu-PAGE gel; (E) same for IBV E and mutant. Results in presence of DTT were similar and are not shown.

shown), although in this case MALDI-TOF/MS mass spectra produced no multimers (Fig. 4A and B). It is thus likely that the presence of multimers in MALDI mass spectra is related to the presence of inter-molecular disulfide bonds.

Expression and purification of IBV E

The fusion construct formed by 6His-BBP$_{\Delta M}$ and IBV E (Fig. 2A and B), also accumulated in inclusion bodies. The construct was eluted from a Ni-NTA column with 250 mM imidazole (Fig. 2D, lane E1) and the yield after this step was similar to that of SARS-CoV E. The HPLC chromatogram of this sample after CNBr cleavage (Fig. 3B) shows two main peaks. Peak 1 corresponds to BBP$_{\Delta M}$ (not shown), whereas the MALDI-TOF/MS spectrum corresponding to peak 2 produced a peak at 12,196.4 Da corresponding to IBV E (calculated 12,116 Da). In this case, a peak at 6,075.6 Da, resulting from double protonation, was observed. The cysteine less mutant also could be purified (Fig. 4C). The identity of IBV E was confirmed by tryptic digestion (see Supplementary File, Fig. S3).

SDS–PAGE electrophoresis of SARS-CoV E and IBV E proteins

The fraction corresponding to peak 2 in the HPLC chromatograms (Fig. 3, inserts) and the equivalent peaks for other similar chromatograms was analyzed by SDS electrophoresis. Both wild type SARS-CoV E and IBV E contain cysteines, therefore the gels were run with or without reductant (DTT). The monomer of wild type SARS-CoV E migrated with its expected mobility in 4–12% Nu–PAGE (~8 kDa) together with several higher oligomers, and mutants SSS and AAA produced only monomers (Fig. 5A). In the presence of DTT (Fig. 5B), wild type SARS-CoV E produced only two bands consistent with monomers and trimers, whereas the mutants were unaffected. The difference between triple mutants and wild type in presence of DTT suggests that cysteine side chains are involved in SARS-CoV E inter-monomeric contacts, but this involvement does not necessarily imply the formation of disulfide bonds. The side chain of cysteine is more hydrophobic [46] than alanine or serine, and the latter residues may lead to different secondary structure propensities relative to cysteine [23,24].

Wild type IBV E produced only one main band in SDS (Fig. 5C), consistent with monomers, and only very faint bands corresponding to higher oligomers. Similar bands were observed for the double cysteine mutant AA, with or without DTT, which indicates that IBV E as prepared did not contain significant amount of inter-monomeric disulfide bonds.

Oligomerization of SARS-CoV E and IBV E in PFO

To confirm that the proteins purified are able to form pentameric oligomers, we used perfluorooctanoic acid (PFO) electrophoresis. PFO is a detergent known to facilitate native interactions between transmembrane α-helices [47,48]. We have shown previously [19] that the synthetic transmembrane domain of SARS-CoV E (SARS-CoV E-TM) forms pentamers in this detergent. However, neither

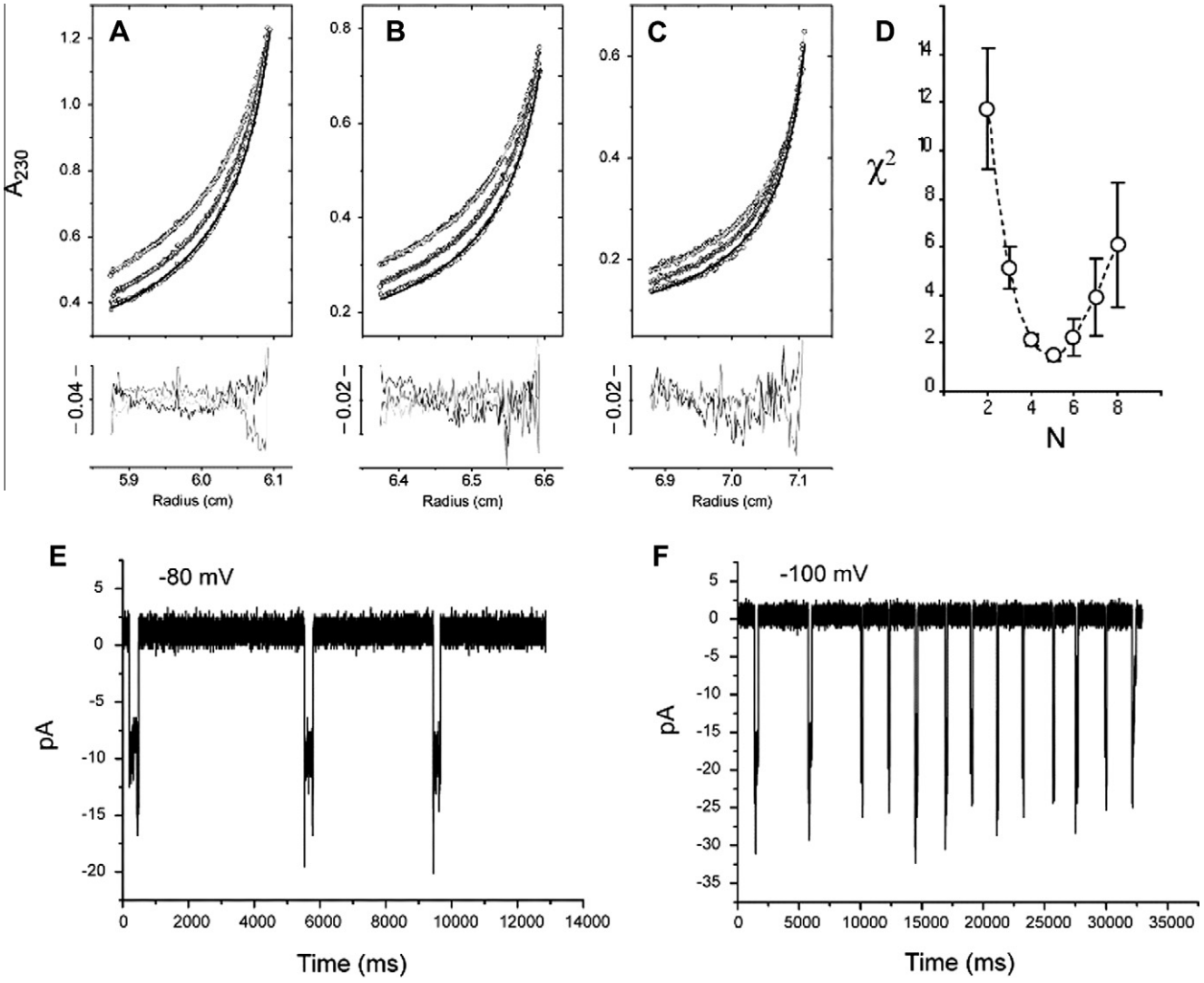

Fig. 6. Sedimentation equilibrium results of SARS-CoV E in C14 betaine micelles. (A–C) SARS-CoV E in the presence of TCEP at 48.000 rpm (black), 40.000 rpm (dark grey), 30.000 rpm (light grey). Residuals are shown at the bottom of each graph. The error (global red chi-squared) was always lower than 1.5; (D) Chi-squared value as a function of the oligomeric size (*N*) in a monomer:*N*-mer self-association model for SARS-CoV E in the presence of TCEP; (E and F) channel activity of purified full length SARS–CoV E in black lipid membranes. Current traces recorded at −80 mV (E) and −100 mV (F).

full length wild type SARS-CoV E or the SSS mutant showed any band consistent with pentamers (Fig. 5D), and produced instead aggregation well above 100 kDa. The AAA mutant, in contrast, showed a clear band consistent with pentamers, ~40 kDa (Fig. 6D). The difference in behavior of this mutant may be due to a change in secondary structure, possibly increased helicity, induced by substitution of Cys to Ala [50]. The presence of DTT did not have any effect on mobility in PFO gels (not shown). In the case of IBV E (Fig. 5E), both for wild type and AA mutant, a large smear was observed around 66 kDa, which is consistent with a predominantly pentameric form.

Analytical ultracentrifugation and channel activity of SARS-CoV E

To further confirm the presence of a pentameric arrangement in these coronavirus envelope proteins, we used a milder, zwitterionic detergent, C14 betaine, in a sedimentation equilibrium (SE) experiment. For wild type full length SARS-CoV E no simple reversible association 'monomer:*N*-mer' model could be fitted to the data (*N* was tested from 2 to 15), probably because of the presence of disulfide bonds. However, in the presence of reductant TCEP, the data could be best fitted to a simple monomer-pentamer self-association model (Fig. 6A–D), with $\log K_a = 17.5 \pm 0.2$, similar to that obtained previously in similar conditions [19] for synthetic peptide TM_{8-37} in similar conditions. This suggests that the main contribution to pentamerization in SARS-CoV E is provided by the TM domain. The fact that we could fit the data in presence of TCEP to a monomer pentamer model demonstrates that disulfide bonds are not required to form pentamers. When we used the cysteine-less mutants (SSS or AAA) or IBV E and mutant AA, the data could still be fitted to a reversible monomer-pentamer equilibrium (not shown). For IBV E a good fit of the data could be obtained without the need of TCEP, consistent with the observation that no disulfide bonds were observed in SDS (Fig. 5C). Purified SARS-CoV E has channel activity in lipid bilayers, although the film was unstable and recordings could only be obtained at two potentials (Fig. 6E and F).

Structural integrity of SARS-CoV E protein in DPC detergent micelles and in lipid bilayers

After confirming that SARS-CoV E is pentameric and functinal for channel activity, we tested with independent techniques if it can be suitable for structural studies and can be properly incorporated in

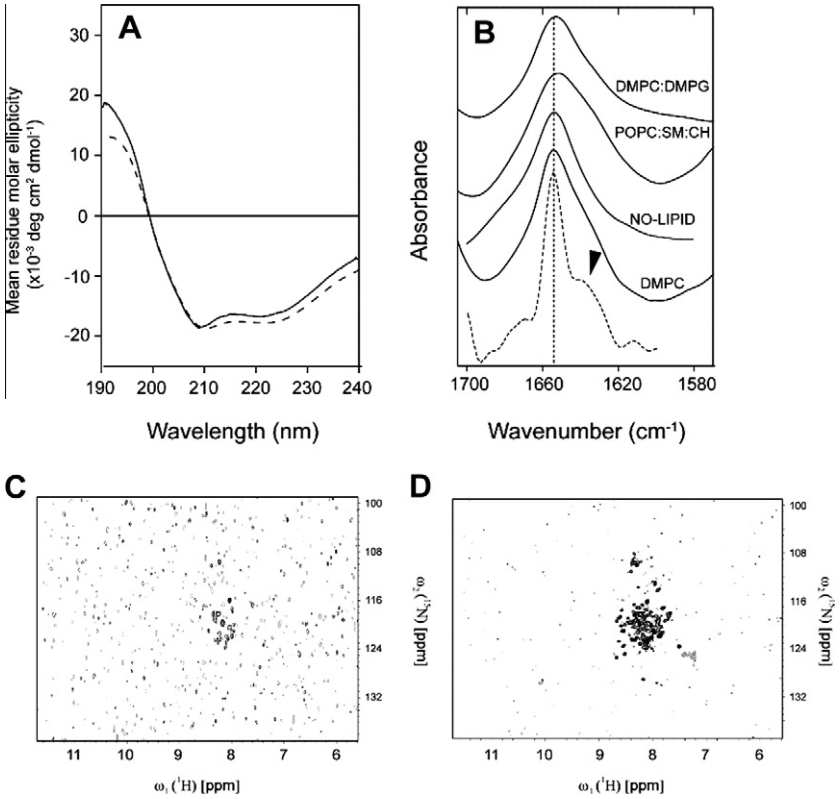

Fig. 7. Structural integrity of purified SARS-CoV E in detergent micelles and lipid bilayers. (A) Circular dichroism spectra of SARS-CoV E in detergent micelles formed by SDS (solid line) or DPC (broken line); (B) Transmission amide I infrared spectra of SARS-CoV E incorporated in multilamellar liposomes of the composition indicated, in D_2O solution. Sample labeled 'no lipid' refers to SARS-CoV E dry after being exposed to HFIP solvent. The deconvolved spectrum corresponding to the sample in DMPC is shown at the bottom (broken line), with the small -structure contribution indicated by an arrow; (C) 2D ^{1}H–^{15}N HSQC spectra for ^{15}N-labeled SARS-CoV E wild type protein in presence of 100 mM DPC and 40 mM DTT at 328 K; (D) same conditions as in (C) but after addition of 100 mM SDS (see details in Materials and Methods).

membranes. Thus we obtained CD and FTIR spectra to assess the secondary structure of SARS-CoV E. The CD spectrum of SARS-CoV E protein incorporated in anionic SDS or lipid-like DPC detergent micelles showed characteristic minima at 208 nm and 222 nm (Fig. 7A), typical of α-helical conformation. This is confirmed by the infrared amide I band envelope (Fig. 7B) obtained after reconstitution of SARS-CoV E protein in lipidic membranes, which shows a predominant band centered at ~1655 cm^{-1}, also characteristic of α-helices, although in some cases (e.g., DMPC) a small shoulder was observed consistent with some β-structure.

Finally, 2D-HSQC NMR spectra were collected in DPC detergent (Fig. 7C) or a mixture of DPC/SDS (Fig. 7D). Although the spectra are not yet of sufficient quality for complete structure determination, the best spectra were obtained using 0.5 mM SARS-CoV E, 100 mM DPC, 100 mM SDS and 40 mM DTT, at a temperature of 328 K (Fig. 7D). From visual inspection, about 70% of the ^{1}H–^{15}N peaks can be detected. Using the triple cysteine mutant did not improve the quality of the spectra (not shown). Nevertheless, the NMR spectrum, combined with CD and FTIR spectra, show that the protein is folded and after further optimization to improve signal intensity and peak dispersion, structural determination should be possible.

We have also shown that these proteins form pentameric oligomers, are functional and are likely be properly folded when incorporated in detergent micelles or in lipid bilayers. We stress the fact that (1) this is the first report that shows unequivocally that coronavirus envelope proteins (full length) are pentameric, and (2) that the mass spectroscopy determination of these proteins is by no means routine. Indeed, the fluctuation between measurements observed for SARS-CoV E was not observed previously in TM fragments of SARS-CoV E [18,19], in proteolytic fragments after tryptic digestion (Supplementary File), or in similar viroporins, e.g., small hydrophobic (SH) protein from RSV [51], which do not form large aggregates in SDS and PFO gels. Therefore, we attribute this variability to the high aggregation tendencies of these proteins.

Overall, these results show that the two coronavirus E proteins could be expressed and purified using a novel fusion protein consisting of a modified transmembrane β-barrel. This is an approach that can be used to study the structural variations and functional characteristics of this so far unexplored family of membrane proteins. This approach can also be used in other small membrane proteins, as long as they do not contain internal methionines.

Acknowledgement

J.T. thanks the National Research Foundation (NRF) Grant NRF-CRP4-2008-2 for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pep.2012.07.005.

References

[1] J.M. Gonzalez, P. Gomez-Puertas, D. Cavanagh, A.E. Gorbalenya, L. Enjuanes, A comparative sequence analysis to revise the current taxonomy of the family *Coronaviridae*, Arch. Virol. 148 (2003) 2207–2235.
[2] S.G. Siddell, The *Coronaviridae*; An Introduction, Plenum Press, New York, NY, 1995.
[3] L. Enjuanes, D. Brian, D. Cavanagh, K. Holmes, M.M.C. Lai, H. Laude, P. Masters, P. Rottier, S.G. Siddell, W.J.M. Spaan, F. Taguchi, P. Talbot, *Coronaviridae*, in:

van Regenmortel, C.M. Fauquet, D.H.L. Bishop, E.B. Carsten, M.K. Estes, S.M. Lemon, D.J. McGeoch, J. Maniloff, M.A. Mayo, C.R. Pringle, R.B. Wickner (Eds.), Virus Taxonomy. Classification and Nomenclature of Viruses, Academic Press San Diego, pp. 835–849.
[4] A.E. Gorbalenya, E.J. Snijder, W.J. Spaan, Severe acute respiratory syndrome coronavirus phylogeny: toward consensus, J. Virol. 78 (2004) 7863–7866.
[5] L. Kall, A. Krogh, E.L. Sonnhammer, Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server, Nucleic Acids Res. 35 (2007) W429–W432.
[6] M.L. Dediego, L. Pewe, E. Alvarez, M.T. Rejas, S. Perlman, L. Enjuanes, Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice, Virology 376 (2008) 379–389.
[7] E.C. Bos, W. Luytjes, H.V. van der Meulen, H.K. Koerten, W.J. Spaan, The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus, Virology 218 (1996) 52–60.
[8] E. Corse, C.E. Machamer, Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles, J. Virol. 74 (2000) 4319–4326.
[9] H. Vennema, G.J. Godeke, J.W. Rossen, W.F. Voorhout, M.C. Horzinek, D.J. Opstelten, P.J. Rottier, Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes, EMBO J. 15 (1996) 2020–2028.
[10] P. Baudoux, C. Carrat, L. Besnardeau, B. Charley, H. Laude, Coronavirus pseudo particles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes, J. Virol. 72 (1998) 8636–8643.
[11] E. Corse, C.E. Machamer, The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction, Virology 312 (2003) 25–34.
[12] E. Mortola, P. Roy, Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system, FEBS Lett. 576 (2004) 174–178.
[13] F. Fischer, C.F. Stegen, P.S. Masters, W.A. Samsonoff, Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly, J. Virol. 72 (1998) 7885–7894.
[14] K.M. Curtis, B. Yount, R.S. Baric, Heterologous gene expression from transmissible gastroenteritis virus replicon particles, J. Virol. 76 (2002) 1422–1434.
[15] J. Ortego, D. Escors, H. Laude, L. Enjuanes, Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome, J. Virol. 76 (2002) 11518–11529.
[16] L. Kuo, K.R. Hurst, P.S. Masters, Exceptional flexibility in the sequence requirements for coronavirus small envelope protein function, J. Virol. 81 (2007) 2249–2262.
[17] Y. Liao, J. Lescar, J.P. Tam, D.X. Liu, Expression of SARS-coronavirus envelope protein in *Escherichia coli* cells alters membrane permeability, Biochem. Biophys. Res. Commun. 325 (2004) 374–380.
[18] K. Pervushin, E. Tan, K. Parthasarathy, L. Xin, F.L. Jiang, D. Yu, A. Vararattanavech, T.W. Soong, D.X. Liu, J. Torres, Structure and inhibition of the SARS coronavirus envelope protein ion channel, PLoS Pathog. 5 (7) (2009) e1000511, http://dx.doi.org/10.1371/journal.ppat.1000511.
[19] K. Parthasarathy, L. Ng, X. Lin, D.X. Liu, K. Pervushin, X. Gong, J. Torres, Structural flexibility of the pentameric SARS coronavirus envelope protein ion channel, Biophys. J. 95 (2008) L39–L41.
[20] J. Torres, K. Parthasarathy, X. Lin, R. Saravanan, A. Kukol, D.X. Liu, Model of a putative pore: the pentameric α-helical bundle of SARS coronavirus E protein in lipid bilayers, Biophys. J. 91 (2006) 938–947.
[21] Y. Liao, Q. Yuan, J. Torres, J.P. Tam, D.X. Liu, Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein, Virology 349 (2006) 264–275.
[22] Q. Yuan, Y. Liao, J. Torres, J.P. Tam, D.X. Liu, Biochemical evidence for the presence of mixed membrane topologies of the Severe Acute Respiratory Syndrome coronavirus envelope protein expressed in mammalian cells, FEBS Lett. 580 (2006) 3192–3200.
[23] L. Wilson, P. Gage, G. Ewart, Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication, Virology 353 (2006) 294–306.
[24] L. Wilson, C. McKinlay, P. Gage, G. Ewart, SARS coronavirus E protein forms cation-selective ion channels, Virology 330 (2004) 322–331.
[25] A. Kuliopulos, C.T. Walsh, Production, purification, and cleavage of tandem repeats of recombinant peptides, J. Am. Chem. Soc. 116 (1994) 4599–4607.
[26] T.J. Park, J.S. Kim, S.S. Choi, Y. Kim, Cloning, expression, isotope labeling, purification, and characterization of bovine antimicrobial peptide, lactophoricin in *Escherichia coli*, Protein Expr. Purif. 65 (2009) 23–29.
[27] A. Pautsch, G.E. Schulz, Structure of the outer membrane protein A transmembrane domain, Nat. Struct. Biol. 5 (1998) 1013–1017.
[28] M.U. Johansson, S. Alioth, K. Hu, R. Walser, R. Koebnik, K. Pervushin, A minimal transmembrane beta-barrel platform protein studied by nuclear magnetic resonance, Biochemistry 46 (2007) 1128–1140.
[29] C. Diefenderfer, J. Lee, S. Mlyanarski, Y. Guo, K.J. Glover, Reliable expression and purification of highly insoluble transmembrane domains, Anal. Biochem. 384 (2009) 274–278.
[30] E. Gordon, R. Horsefield, H.G. Swarts, J.J. de Pont, R. Neutze, A. Snijder, Effective high-throughput overproduction of membrane proteins in *Escherichia coli*, Protein Expr. Purif. 62 (2008) 1–8.
[31] E. Gross, B. Witkop, Nonenzymatic cleavage of peptide bonds: the methionine residues in bovine pancreatic ribonuclease, J. Biol. Chem. 237 (1962) 1856–1860.
[32] E. Corse, C.E. Machamer, The cytoplasmic tail of infectious bronchitis virus E protein directs Golgi targeting, J. Virol. 76 (2002) 1273–1284.
[33] L.A. Lopez, A.J. Riffle, S.L. Pike, D. Gardner, B.G. Hogue, Importance of conserved cysteine residues in the coronavirus envelope protein, J. Virol. 82 (2008) 3000–3010.
[34] J.A. Boscarino, H.L. Logan, J.J. Lacny, T.M. Gallagher, Envelope protein palmitoylations are crucial for murine coronavirus assembly, J. Virol. 82 (2008) 2989–2999.
[35] E.B. Thorp, J.A. Boscarino, H.L. Logan, J.T. Goletz, T.M. Gallagher, Palmitoylations on murine coronavirus spike proteins are essential for virion assembly and infectivity, J. Virol. 80 (2006) 1280–1289.
[36] I. Bertini, I. Gelis, N. Katsaros, C. Luchinat, A. Provenzani, Tuning the affinity for lanthanides of calcium binding proteins, Biochemistry 42 (2003) 8011–8021.
[37] Y. Ye, H.W. Lee, W. Yang, J.J. Yang, Calcium and lanthanide affinity of the EF-loops from the C-terminal domain of calmodulin, J. Inorg. Biochem. 99 (2005) 1376–1383.
[38] C.M. Bauer, L.H. Pinto, T.A. Cross, R.A. Lamb, The influenza virus M2 ion channel protein: probing the structure of the transmembrane domain in intact cells by using engineered disulfide cross-linking, Virology 254 (1999) 196–209.
[39] A. Shevchenko, H. Tomas, J. Havlis, J.V. Olsen, M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, Nat. Protoc. 1 (2006) 2856–2860.
[40] W. Meng, H. Zhang, T. Guo, C. Pandey, Y. Zhu, O.L. Kon, S.K. Sze, One-step procedure for peptide extraction from in-gel digestion sample for mass spectrometric analysis, Anal. Chem. 80 (2008) 9797–9805.
[41] G.G. Kochendoerfer, D. Salom, J.D. Lear, R. Wilk-Orescan, S.B. Kent, W.F. DeGrado, Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly 1, Biochemistry 38 (1999) 11905–11913.
[42] R. Li, C.R. Babu, J.D. Lear, A.J. Wand, J.S. Bennett, W.F. DeGrado, Oligomerization of the integrin αIIbβ3: roles of the transmembrane and cytoplasmic domains, Proc. Nat. Acad. Sci. USA 98 (2001) 12462–12467.
[43] P. Schuck, On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation, Anal. Biochem. 320 (2003) 104–124.
[44] T.M. Laue, W.F. Stafford 3rd, Modern applications of analytical ultracentrifugation, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 75–100.
[45] N. Nagano, M. Ota, K. Nishikawa, Strong hydrophobic nature of cysteine residues in proteins, FEBS Lett. 458 (1999) 69–71.
[46] J. Maupetit, P. Derreumaux, P. Tuffery, PEP-FOLD: an online resource for de novo peptide structure prediction, Nucleic Acids Res. 37 (2009) W498–W503.
[47] M. Ramjeesingh, L.J. Huan, E. Garami, C.E. Bear, Novel method for evaluation of the oligomeric structure of membrane proteins, Biochem. J. 342 (Pt 1) (1999) 119–123.
[48] A.G. Therien, C.M. Deber, Oligomerization of a peptide derived from the transmembrane region of the sodium pump gamma subunit: effect of the pathological mutation G41R, J. Mol. Biol. 322 (2002). 583-550.
[50] C. Combet, C. Blanchet, C. Geourjon, G. Deleage, NPS@: network protein sequence analysis, Trends Biochem. Sci. 25 (2000) 147–150.
[51] S.W. Gan, E. Tan, X. Lin, D. Yu, J. Wang, G. Tan, A. Vararattanavech, C.Y. Yeo, C.H. Soon, T.W. Soong, K. Pervushin, J. Torres, The small hydrophobic protein of the human respiratory syncytial virus forms pentameric ion channels, J. Biol. Chem. 287 (29) (2012) 24671–24689 (Epub 23 May 2012).