Evolutionarily diverse caveolins share a common structural framework built around amphipathic discs

Bing Han [1,2,†], Sarah Connolly [3,†], Louis F. L. Wilson [2,4], Darrin T. Schultz [5], Alican Gulsevin [6,‡] Jens Meiler [6,7], Erkan Karakas [8], Melanie D. Ohi [3,9,*], and Anne K. Kenworthy [1,2,*]

[1] Center for Membrane and Cell Physiology, University of Virginia, Charlottesville, VA USA.
[2] Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA, USA
[3] Life Sciences Institute, University of Michigan, MI, USA
[4] Howard Hughes Medical Institute, University of Virginia School of Medicine, Charlottesville, VA, USA
[5] Department of Neuroscience and Developmental Biology, University of Vienna, Vienna 1010, Austria
[6] Department of Chemistry, Vanderbilt University Nashville, TN, USA
[7] Institute for Drug Discovery, Leipzig University, Germany
[8] Department of Molecular Physiology and Biophysics, Vanderbilt University Nashville, TN, USA
[9] Department of Cell and Developmental Biology, University of Michigan, MI, USA

[†] These authors contributed equally to this work.

[‡] Current address: Department of Pharmaceutical Sciences, College of Pharmacy and Health Sciences, Butler University, Indianapolis, IN

* Correspondence: akk7hp@virginia.edu, mohi@umich.edu

Short title: Structures of caveolins across evolution

Teaser: Structures of evolutionarily distant caveolins reveal an ancient, conserved lipid-protein interface predating metazoans.

Abstract

Caveolins represent a unique family of membrane-remodeling proteins widely expressed across animals (Metazoa). In vertebrates, caveolins are integral to plasma membrane homeostasis, forming flask-shaped invaginations known as caveolae. Recent studies demonstrate that human caveolin-1 assembles into an amphipathic disc composed of 11 spirally packed protomers. However, the structural basis underlying caveolin function across diverse organisms remains elusive. In this study, we computationally predicted structures for 74 evolutionarily diverse caveolins, including a newly identified choanoflagellate caveolin, identifying seven conserved structural elements and a propensity for amphipathic disc formation. Despite extreme sequence variability, cryo-EM structures of caveolins from the purple sea urchin *Strongylocentrotus purpuratus* and the choanoflagellate *Salpingoeca rosetta* exhibit striking structural similarities to human caveolin-1, strengthening the computational predictions. These results suggest that caveolins share an ancient structural framework that predates Metazoa and provide a new structure-based paradigm to explore the molecular basis of caveolin function across evolutionary space.

Introduction

Eukaryotes contain elaborate endomembrane systems consisting of morphologically and functionally distinct membrane-bound compartments. The construction and maintenance of these compartments relies on the actions of ancient families of proteins capable of remodeling membranes (*1, 2*). To support enhanced requirements for cell-cell adhesion, communication, and signaling during the transition from single-celled eukaryotes to animals (Metazoa), a dramatic expansion in membrane-associated proteins occurred (*1*). Among the membrane proteins thought to have first emerged in Metazoa is the caveolin family of membrane remodeling proteins (*1-3*). Best recognized for their role in vertebrates as structural components of flask-shaped invaginations of the plasma membrane known as caveolae, caveolins homologues have been identified in most metazoan clades (*2-4*), suggesting they have fulfilled essential biological roles since the ancestor of the clade existed approximately 800 million years ago (*5*). In humans, caveolins and caveolae are distributed throughout the body and serve as important regulators of multiple organ systems (*6-9*). Furthermore, caveolins and caveolae have been broadly implicated in regulation of cell signaling, lipid metabolism, and sensing and responding to stress (*8, 10-13*).

Unlike traditional vesicle coat proteins such as clathrin, COPI, and COPII that cycle on and off membranes and share evolutionary origins and structural features (*14*), caveolins are unrelated in sequence to other vesicle coat proteins and remain integrated in membranes throughout their life cycle. Most studies of caveolins have focused on their roles in mammalian cells where caveolae are often abundant (*15-17*). However, in some cell types lacking cavins, a second family of proteins required for caveolae biogenesis, caveolins can function independently of caveolae (*18-21*). Since the cavin family appears to be found only in vertebrates (*22*), this suggests that in most organisms caveolins function independently of classically defined caveolae. To date, however, only a handful of examples of caveolins from non-vertebrate organisms have been studied (*3, 23-26*). Thus, our knowledge of the functions of caveolins across evolutionary space is currently extremely limited.

Until recently, the molecular architecture of caveolins was unknown. A breakthrough in our understanding of caveolins emerged with the discovery of the structure of a homo-oligomeric 8S complex of human caveolin-1 (Cav1), an essential component of caveolae in non-muscle cells, using cryo-electron microscopy (cryo-EM) (*27, 28*). The complex is composed of 11 Cav1 protomers symmetrically arranged into an unusual amphipathic disc-like structure predicted to fully insert into the cytoplasmic leaflet of the plasma membrane (*27-29*). However, whether

related proteins from other organisms (including distantly related caveolin homologues) behave in the same way is completely unknown.

Here, we report that in addition to being expressed in metazoans, caveolin homologs exist in choanoflagellates, suggesting they are of pre-metazoan origin. Using a combination of computational, phylogenetic, and structural approaches, we show that despite extreme sequence variability, even the most distantly related caveolin homologues share a surprisingly well-conserved set of structural elements. These findings suggest that caveolins share an ancient and conserved structural framework that diverse organisms have co-opted to fulfill distinct physiological roles. They also provide a new framework to probe the structural basis for the function of caveolins across evolution.

Results

Consistent evidence for the existence of choanoflagellate caveolins

Caveolins have typically been treated as a metazoan-specific family (*1-3, 30*). However, one report based on BLAST (basic local alignment search tool) and PFAM (protein family) searches hinted at the presence of a potential caveolin homologue in the genome of *Monosiga ovatus* (*31*), a member of the choanoflagellates (a group of unicellular and colonial organisms constituting the closest sister to animals). Unfortunately, this reported sequence is not publicly available for analysis. Nevertheless, one further sequence from the choanoflagellate *Salpingaeca rosetta* (strain: ATCC 50818) is currently annotated in the UniProt database as a caveolin (UniProt: F2U793). Surprisingly, this sequence shares almost no recognizable sequence identity with human Cav1 (13%); however, if truly related, it would represent the most evolutionarily divergent caveolin found to date. Thus, we were interested to verify this annotation by attempting to reproduce this hit independently. To this end, we built a caveolin HMM (hidden Markov model) profile from an alignment based on a previously identified region of high conservation (*3*). Encouragingly, HMMER searches against the *S. rosetta* proteomes retrieved the F2U793 sequences with an E-value of 8.8×10^{-11}, indicating a confident prediction of the homology. By contrast, a sequence from the filasterean *Capsaspora owczarzaki* (strain: ATCC 30864) annotated in UniProt as a caveolin (A0A0D2VH37) failed to appear as a significant hit when the HMM profile was searched against the *C. owczarzaki* genome, suggesting this annotation is spurious. These findings support the idea that caveolins are in fact not limited to Metazoa but are also found in some choanoflagellates, the closest living relatives of Metazoa.

Caveolin protomers are predicted to organize into disc-shaped complexes composed of spiraling amphipathic α-helices.

The existence of such distantly related caveolin homologues with limited sequence similarity raises the question as to whether they have similar folds and/or functions with human Cav1. The cryo-EM structure of human Cav1 showed that the protein assembles into an amphipathic disc composed of 11 protomers (*27*). N-terminal sequences located on the outer rim of the disc consist of a pin motif and oligomerization domain that play a critical structural role in interlocking adjacent protomers together to form a flat disc of spiraling protomers. Moving from the N-terminal domains towards the C-terminus, each protomer consists of a central curved amphipathic α-helix that ends in a β-strand formed by the C-terminal residues. In the context of the complex, the α-helix from each protomer spirals toward the center of the disc and the eleven C-terminal β-strands form a parallel β-barrel. Thus, the N-terminal rim regions, the central α-helices, and C-terminal β-strands are all important for the stability of the complex.

To investigate whether other metazoan caveolins and the newly identified choanoflagellate caveolin share similar structural features, we used AlphaFold2 (AF2) as a predictive tool (*32*).

Despite the unusual features of human Cav1, AF2 is able to predict its overall fold and reproduce its ability to oligomerize to form an amphipathic, disc-shaped structure characteristic of the Cav1 8S complex (*33*). This suggests AF2 is a useful tool to probe the structural properties of the caveolin family. Using AF2, we predicted the structures of Pfam-annotated caveolin family members from representative species of 14 different metazoan phyla/superphyla and the choanoflagellate *S. rosetta* (**Fig 1**).

For our initial analysis, we generated *n*-mers of increasing size with AF2.1 (**Supplementary File 1**) and then generated the 11-mer with AF2.2 (**Supplementary File 2**). The results of this analysis predict that caveolins in essentially all the examined species form closely packed discs or rings that spiral in a clockwise direction when viewed from the cytoplasmic face (**Fig 1**). Similar to the structure of human Cav1, most of the caveolins were predicted to contain N-terminal disordered regions located around the outer rim of the complex and central β-barrels formed by parallel β-strands. Interestingly, C-terminal disordered regions emanating from the central β-barrel were also visible in several caveolins (**Fig 1**). Together, these results suggest that even in the most distantly related species, caveolins can homo-oligomerize and contain similar structural elements as human Cav1. Furthermore, they further support the idea that the *S. rosetta* protein could be homologous to vertebrate caveolins. In contrast, AF2's prediction for the sequence A0A0D2VH37 from *Capsaspora owczarzaki* does not support it being a true caveolin (**Supplementary File 2**). This is consistent with the conclusions drawn from the sequence alignment.

The disc-shaped human Cav1 complex has a hydrophobic membrane facing surface and outer rim (*27*). To determine if other caveolins share this feature, we compared the distribution of hydrophobic and hydrophilic residues on the predicted structures. In all models examined, the long α-helices forming the disc are predicted to be amphipathic, consisting of a hydrophobic membrane-interacting surface and a hydrophilic surface that would face the cytoplasm (**Fig 1**). The regions of the proteins forming the outer rims of the complexes are also amphipathic. Additionally, the C-terminal β-strand, in the proteins predicted to have this feature, are also predicted to be amphipathic, making the interior of resulting β-barrels hydrophobic and the exterior hydrophilic. These analyses predict that a wide range of caveolins likely assemble into amphipathic discs generated by an arrangement of protomers with spiraling α-helices.

Caveolins consist of seven basic structural units

We next sought to identify common structural motifs shared by these diverse caveolins. Previous studies identified a series of functionally important domains across mammalian caveolins, including the oligomerization, scaffolding, and intramembrane domains (*34-38*). These domains were primarily identified by sequence analysis or truncation studies and do not map in a straightforward way to the experimentally determined structure of human Cav1 (*27*). However, it is now possible to define specific molecular features of caveolin using the structure of human Cav1 as a template. To illustrate their position in the structure and relationships to one another, we mapped these conserved structural motifs onto the experimentally determined structure of a Cav1 protomer built from the cryo-EM density map and onto the computationally predicted structures of human Cav1 and *C. elegans* caveolin (Q94051) (**Fig 2**). Moving from the N- to C-terminus of the proteins, we identified seven basic structural units:

1) N-terminal variable region. In human Cav1, residues 1-48 are predicted to be disordered and were not resolved in the cryo-EM structure (*27*). Similarly, other caveolins are predicted to contain N-terminal disordered regions (**Fig 1**). For *C. elegans* caveolin Q94051, our model predicts an unstructured and flexible N-terminal region consisting of 99 residues (**Fig 2, yellow**).

2) Pin motif. First identified in the experimental structure of the human Cav1 8S complex (*27*), the pin motif (residues 49-60 in human Cav1) makes critical contacts with each neighboring

protomer at the rim of the 8S complex (*27*). A similar motif is predicted to exist in *C. elegans* caveolin Q94051 (**Fig 2, red**).

3) Hook structure. Residues 61-81 of human Cav1 consist of a loop that undergoes a 180° turn (**Fig 2, blue**). A similar hook-shaped structural motif is predicted to exist in *C. elegans* caveolin Q94051 (**Fig 2**). This structural element corresponds to the first half of the oligomerization domain (residues 61-101) of human Cav1. Embedded within this same region is the highly conserved signature motif (residues 68-75), consisting of a 3/10 helix followed by a short-coiled structure.

4) Scaffolding domain. The region encompassing residues 82-101 of human Cav1 is traditionally identified as the caveolin scaffolding domain (CSD). This segment corresponds to the initial α-helix (α-1) of the Cav1 protomer, as revealed by the cryo-EM structure. Within the 8S complex, the α-1 helices form the periphery of the disc. In relation to the location of the detergent micelle in the cryo-EM map (*27*), the N-terminal residues of the α-1 helix would face the cytoplasm, while the central and C-terminal regions of the α-1 helix would interface with the aqueous lipid bilayer boundary and the hydrophobic core of the membrane, respectively. Importantly, the α-1 helix extends beyond the classical boundaries of the CSD. Thus, we redefined the entire α-1 region (residues 81-107) as the scaffolding domain, considering its cohesive functional role in the experimental and predicted structures (**Fig 2**).

5) Spoke region. Residues 108-169 of human Cav1 consist of a series of α-helices connected in tandem by short loops. This generates a long helical region in the middle of the protein with all the turns and kinks occurring along the same orientation and plane forming a semi-circle arc of about 180°. We now define the spoke region to encompass residues previously referred to as the intramembrane domain (IMD) (residues 102-134) as well as the helical region we previously designated as the spoke-like region (residues 135-169) since they are not structural distinct (**Fig 2, gray**). The presence of a "bow-shaped" spoke region is a conserved structural feature in most of the caveolins we examined (**Supplemental Files 1 and 2**).

6) C-terminal β-strand. Residues 170-176 of the C-terminal domain of the human Cav1 protomer folds into a β-strand that assembles into an amphipathic parallel β-barrel together with neighboring protomers (**Fig 2, cyan**). Many other caveolins are also predicted to contain C-terminal β-strands (**Fig 1, Supplemental Files 1 and 2**).

7) C-terminal variable region. While the structure of human Cav1 essentially ends in a β-strand, a subset of other caveolins contain an additional C-terminal region that differs in length and composition across caveolins (**Fig 2, purple**). Accordingly, we refer to this region as the C-terminal variable domain. The structure of this region is typically predicted by AF2 with low confidence, suggesting it is disordered (**Supplemental Files 1 and 2**).

To investigate how these structural elements are utilized by different caveolins and how they change during evolution, we next assessed their phylogenic relationships.

Phylogenetic analysis reveals the existence of a non-metazoan caveolin and a previously unidentified clade(s)

As discussed above, computational modeling predicted that the structure of the caveolin protein from *S. rosetta* is similar to that of human Cav1, suggesting it could be homologous to vertebrate caveolins. However, its specific evolutionary relationship with metazoan caveolins is unknown. In addition, many new metazoan species have been sequenced since the last comprehensive phylogenetic analysis of the caveolin family was performed (*3*). We hence set out to conduct an updated phylogenetic analysis of the caveolin family. To do so, we inferred a maximum-likelihood phylogeny using a set of previously categorized caveolin and caveolin-like sequences (*3*), together with the 74 protein sequences with caveolin Pfam annotations from 15 distantly related

holozoans (**Table S1**). Aside from *S. rosetta*, *Amphimedon queenslandica* (Porifera, Metazoa) is of evolutionary interest due to its early divergence with respect to vertebrates (*39, 40*). With the goal of characterizing caveolin evolution across animal diversity, we expanded the search to identify additional caveolin sequences from the two clades most distantly related to all other animals (*41*): sponges (Porifera, Metazoa) and comb jellies (Ctenophora, Metazoa).

In the *A. queenslandica* reference proteome we found five proteins with caveolin Pfam annotations (UniProt: A0A1X7UHP5, A0A1X7UGA1, A0A1X7TMH4, A0A1X7VPY7, A0A1X7VRV8). As with the *S. rosetta* sequence, the *A. queenslandica* sequences were highly dissimilar to human CAV1 (11–19% identity). To help verify their homology, we employed a HMMER search against the *A. queenslandica* genome using the same HMM profile as above. As with the *S. rosetta* sequence, *A. queenslandica* hits were retrieved with significant E-values (ranging from 5.1×10^{-15} to 2.5×10^{-11}), suggesting that the putative sponge caveolins may be true homologs to the caveolins found in bilaterians, a major metazoan clade characterized by bilateral symmetry. However, for ctenophores, the phylum of comb jellies, despite using both protein sequences and HMM profiles to search for caveolins in publicly available genome and transcriptome assemblies (*42*), we found no recognizable or credible orthologs of caveolins.

Using this dataset of caveolin sequences, we inferred a maximum likelihood phylogeny (**Fig S1**). Current evidence suggests all vertebrate caveolins are thought to descend from three ancestral sequences: CavX, CavY, and CavZ. CavX and CavZ occupied adjacent positions in the ancestral genome (*3*). Whereas CavY appears to have been lost in most vertebrates, CavX is proposed to have given rise to Cav1 and Cav3, the 'canonical' caveola-forming family members (*3*). CavZ appears to have given rise to the Cav2 and Cav2R groups, Cav2R having later been lost from placental mammals (*3*). Other clades include 'Group 1' protostome sequences that form a monophyletic group with CavY descendants ('CavY extended'), Protostomia Groups 2 and 3, and a more distantly related group coined the 'caveolin-like' group (*3*). Though achieving only weak bootstrap support values, the resulting tree largely replicated the previously proposed clades, consisting of Cav1/3, Cav2/2R, CavY extended, Protostomia Group 1, Protostomia Group 2, and Cav-like (**Fig S1**). Many of the sequences of interest could be tentatively assigned to one of these previously defined groups (**Fig S1**).

Given the long branch length connecting the *S. rosetta* sequence to the rest of the tree and the known evolutionary relationship between choanolagellates and Metazoa, we designated this sequence as a provisional outgroup, named Choa-CAV (**Fig S1**). Correspondingly, we named all metazoan caveolins as Meta-CAV. Within the Meta-CAV group, the selected caveolin homologs from *A. queenslandica* also could not be placed in any of these pre-established clades. These sequences appeared in a monophyletic group that appears to be a sister to all other metazoan sequences, suggesting an early divergence with respect to the other metazoan sequences. We chose to designate this group as "atypical caveolins" and the remaining Meta-CAV groups as "typical caveolins". We further broke down the typical caveolins into Type I and Type II-CAV. Almost all the relatively well studied caveolins, such as human Cav1, Cav2, and Cav3, as well as *C. elegans* caveolins and *Apis mellifera* caveolin, belong to Type II-CAV. Type I-CAV corresponds to the CAV-like clade identified in a previous study (*3*).

We next asked whether these different groups of caveolins shared similar structural frameworks to human Cav1. To test this, we used computational modeling to predict the structures of the 74 caveolins (**Supplemental File 2**). We then scored each structure by whether it contained the seven basic structural units we defined using the human and *C. elegan*s caveolins (**Fig 3, 4A**). To illustrate key similarities and differences across evolutionary distant caveolins we selected four examples taken from the major classes of caveolins (**Fig 4B-N**). They include 1) human Cav1, a Type II-CAV (**Fig 4C-E**); 2) a Type I-CAV from *Strongylocentrotus purpuratus* (A0A7M7T4C2)

(**Fig 4F-H**); 3) an atypical-caveolin from *A. queenslandica* (A0A1X7UHP5) (**Fig 4I-K**); and 4) the Choa-caveolin from *S. rosetta* (F2U793) (**Fig 4L-N**). All four caveolin classes are predicted to contain an N-terminal variable region, hook structure, scaffolding domain, and spoke region (**Fig 4B-N**). Despite being essential for the formation of human Cav1 complexes, the pin motif is found only in Type II caveolins (**Fig 3, 4A**). The presence of the C-terminal β-strand also varied across clades. β-strands were predicted to be absent from 60% of the Atypical caveolins, including the representative *A. queenslandica* caveolin (**Figs 3, 4**), but were predicted to exist in an extended form in other clades. The occurrence of the C-terminal variable region was predicted to vary across caveolins, even within the same clade (**Fig 3**). Finally, although all the 74 caveolin complexes we examined form a disc with a hydrophobic face, approximately half of the complexes, including *S. rosetta* and *A. queenslandica* caveolins, have no charged residues on this face, whereas others, including human Cav1, contain a few charged residues that because of the symmetry of the complex form a charged ring across the hydrophobic face (**Fig 3, Fig S2**).

Together, our computational models predict that evolutionarily distant caveolins share a similar structural organization that includes an N-terminal variable region, a hook structure, a scaffolding domain, and a spoke region. The structural units that appear to be less conserved across evolution, and sometimes even within specific clades, were the pin motif, β-barrel, and C-terminal variable region. However, even in the absence of these structural features, most of the caveolins were still predicted to form amphiphilic disc shaped oligomers.

Electron microscopy (EM) shows diverse caveolins can form disc-shaped oligomers

Computational modeling is useful to generate hypotheses but requires experimental validation (*43*). To test key predictions from our evolutionary and computational modeling analyses, we used negative stain EM to visualize the structure of proteins from the four major classes of caveolins. We expressed human Cav1, *S. purpuratus* (A0A7M7T4C2), *A. queenslandica* (A0A1X7UHP5), and *S. rosetta* (F2U793) sequences in *Escherichia coli* and purified them in detergent using size exclusion chromatography (SEC). SEC traces and western blotting confirmed that these caveolins form high molecular weight complexes (**Fig S3**). To visualize their overall structure, we performed negative stain EM of individual complexes (**Fig 5**).

Negative stain 2D class averages of the complexes from each purification were roughly the same size and most assumed a disc-like geometry although there was some heterogeneity in the shape of the discs (**Fig 5, Fig S4**). *S. purpuratus* and *S. rosetta* caveolins formed uniform discs with distinct inner and outer rings, similar to the appearance of human Cav1 in negative stain (**Fig 5A, B, D**). The *A. queenslandica* caveolin complex, although not as well-ordered, still appeared to form disc-like complexes, but lacked the central density observed in the human, *S. purpuratus,* and *S. rosetta* caveolin class averages (**Fig 5C**). Furthermore, the *S. purpuratus* and *S. rosetta* caveolins migrated as stable complexes on blue native gels (**Fig 5E**). Taken together with our computational models, these results demonstrate that proteins in all four classes of caveolins, even though evolutionarily distant, can assemble into disc shaped complexes, suggesting a conserved structural “fingerprint” for the caveolin family of proteins.

Cryo-EM reveals the molecular basis for the assembly of S. purpuratus caveolin complex

To examine the molecular architecture of the *S. purpuratus* caveolin, we used single particle cryo-EM (**Fig 6**). The same protein sample used for negative stain analysis was vitrified, and images were collected using a Thermo Fisher 300 kV Titan Krios equipped with a Gatan K3 camera and BioQuantum energy filter. An *ab initio* 3D reconstruction with no applied symmetry resulted in a low-resolution density map with 11 spiraling α-helices (**Fig S5**). No other symmetries were detected in extensive 2D or 3D classification analyses. Non-uniform 3D refinement using the *ab initio* model and designating 11-fold symmetry (C11) resulted in a 3.1 Å resolution structure of the

S. purpuratus caveolin complex (**Fig S6, Fig 6A-C**). The resolution of the density map allowed us to build a model of the *S. purpuratus* caveolin spanning amino acids 29-152 (**Fig 6D-F**). Due to a lack of density in the map, we could not build a model for residues 1-28 or 153-159. These N- and C-terminal amino acids are predicted to be disordered (**Fig 3**).

The *S. purpuratus* caveolin complex is ~130 Å in diameter and has a height of ~36 Å (**Fig 6E**). Like the human Cav1 8S complex, the *S. purpuratus* caveolin complex is composed of 11 spiraling α-helical protomers which form a disc-shaped structure (**Fig 6B, E**). The variable N-terminal region extends along the cytoplasmic face of the complex before making a 180˚ turn at the hook structure. Amphipathic α-helices form the spoke region, with the α-1 helix forming the rim of the disc that is ~15 Å in height, and the C-terminus of each protomer forms a β-strand that organizes into a protruding central parallel β-barrel ~30 Å in diameter (**Fig 6G, H**). The hydrophobic face of the disc was surrounded by an unstructured detergent micelle, suggesting that this surface interacts with membrane and the opposite side would face towards the cytoplasm (**Fig 6I**). The hydrophobic interior of the β-barrel also interacted with the detergent micelle (**Fig 6I**), suggesting that it too could interact with lipids; however, no interpretable density was detected in the cryo-EM density map even with no applied symmetry. Direct comparison of the experimental and predicted structures shows AF2.2 predicted much of the *S. purpuratus* caveolin complex secondary structure correctly but over-predicted the length of the β-barrel and underpredicted the curvature of the complex (**Fig 6J-L**).

Cryo-EM structure of the choanoflagellate S. rosetta caveolin complex shows key differences at the N- and C-terminal regions

After determining the structure of the metazoan *S. purpuratus* caveolin complex, we turned to the choanoflagellate caveolin protein from *S. rosetta*. The structure of *S. rosetta'*s caveolin protein and understanding whether it can oligomerize are of particular interest given its evolutionary distance from metazoans as well as its predicted lack of a pin motif.

Using single-particle cryo-EM, we determined the structure of the *S. rosetta* caveolin complex (**Fig 7**). The same protein sample used for negative stain analysis was vitrified, and images were collected using a Thermo Fisher 300 kV Titan Krios equipped with a Gatan K3 camera and a BioQuantum energy filter. 2D classes of the complex in vitrified ice showed that this complex is composed of 11 protomers, and an *ab initio* reconstruction with no applied symmetry also had 11 spiraling helical regions (**Fig S7**). As with our analysis of the human Cav1 and *S. purpuratus* caveolin complexes, neither extensive 2D or 3D classification of the datasets identified any other symmetries. Thus, 11-fold symmetry (C11) was imposed during structural refinement, yielding a cryo-EM density map at a 2.9 Å resolution (**Fig 7A-C**, **Fig S8**). This resolution allowed us to build a model of the *S. rosetta* caveolin protein spanning amino acids 79-231 directly into the density map (**Fig 7D-F**). We were unable to model the predicted disordered N-terminal region (a.a. 1-78) and the most C-terminal residues (a.a. 232-233) due to a lack of density for these regions.

The *S. rosetta* caveolin complex is ~127 Å in diameter and has a height of ~46 Å. This complex has an overall disc-like shape composed of spiraling α-helices with a central β-barrel. The variable N-terminal region first forms a short α-helix that is positioned about halfway up the β-barrel and then continues with a span of residues that snakes along the cytoplasmic facing surface of the complex. The protomer makes a 180˚ turn at the hook structure, and this turn is followed by the spoke region, which consists of a series of curved α-helices. The α-1 helix forms the rim at the edge of the disc that is ~16 Å in height (**Fig 7G-H**). Finally, the *S. rosetta* caveolin protomer has a C-terminal β-strand with a height of ~46 Å which, along with the other 10 β-strands from the other protomers, forms a parallel β-barrel with a diameter of ~30 Å (**Fig 7E**). In the cryo-EM density map one side of *S. rosetta* caveolin complex is surrounded by an unstructured detergent

micelle suggesting that this would face the membrane while the opposite hydrophilic side would face towards the cytoplasm (**Fig 7I**). The detergent micelle also extends into the interior of the hydrophobic β-barrel. The width of the β-barrel is large enough to accommodate lipids; however as with the other caveolin complexes, no interpretable density was detected in the cryo-EM density map even when no symmetry was applied.

Comparison of the predicted and experimental structures reveals that, while AF2.2 accurately predicted that the overall secondary structure of some regions of *S. rosetta* caveolin are similar to human Cav1, it failed to capture the complex organization of the N- and C-terminal portions of *S. rosetta* caveolin or the correct curvature of the disc (**Fig 7J-L**). AF2.2 predicted additional α–helices on the C-terminal region of the complex instead of the extended β-strands observed in the experimental structure, and it predicted the N-terminal α–helices to be oriented perpendicular to the membrane-facing surface as opposed to their parallel placement in the cryo-EM structure. Additionally, the AF2.2 prediction did not capture the concave curvature of the *S. rosetta* caveolin complex (**Fig 7E, K**).

Evolutionarily distant caveolins share structural motifs but differ in dimensions

We next directly compared the structures of the *S. purpuratus* and *S. rosetta* caveolin complexes with the previously determined human Cav1 complex (**Fig 8**). Despite sharing no significant global sequence similarity with human Cav1 (16%/35% identity/similarity for *S. purpuratus* caveolin and 13%/28% for *S. rosetta* caveolin), the secondary structure of the protomers and organization of all three complexes share overall similarities (**Fig 8**). Each complex consists of an 11-mer disc formed by spiraling α-helices and a central β-barrel. The spoke region and scaffolding domain that make up the disc of the complexes contain a similar number of residues for each caveolin (88, 87, and 90 residues for human Cav1, *S. purpuratus* caveolin, and *S. rosetta* caveolin, respectively). However, both the *S. purpuratus* caveolin and *S. rosetta* caveolin complexes are smaller in diameter (~130 Å and ~127 Å, respectively) compared to human Cav1 (~140 Å). Although the *S. purpuratus* and *S. rosetta* caveolin complexes exhibit tighter packing of the η1 and α1 helices than the human Cav1 complex, this packing loosens towards the rim of the complexes. The difference in diameter between the complexes is likely due to the increased curvature of the *S. purpuratus* and *S. rosetta* caveolin complexes compared to the human Cav1 complex. While all three structures have a conserved β-barrel outer diameter of ~28-30 Å, the *S. rosetta* caveolins complex's central β-barrel has a height of ~46 Å, 10 Å longer than the β-barrels seen in the human Cav1 and *S. purpuratus* caveolin complexes (**Fig 8B, E, H**).

Based on our new domain nomenclature, the three caveolins also share similar domain organization, with a few notable variations. *S. purpuratus* caveolin has an unstructured N-terminal variable region not resolved in the cryo-EM density map (a.a.1-28), conserved scaffolding domain (a.a. 55-81), curved spoke region (a.a. 82-142) consisting of a series of connected α-helices, and a C-terminal β-strand (a.a. 143-152) that forms a parallel β-barrel. It is predicted to contain a six-residue C-terminal variable region, which was not resolved in the structure (a.a. 153-159) (**Fig 8, Fig S9**). The *S. rosetta* caveolin protein, even with a more extensively structured N-terminus, still has a long unstructured N-terminal variable region not resolved in the cryo-EM density map (a.a. 1-78). It also has a conserved scaffolding domain (a.a. 125-152), a bow-shaped spoke region (a.a. 153-215), and a C-terminal β-strand (a.a. 216-230) but essentially lacks a C-terminal variable region (**Fig 8, Fig S9**). However, both *S. purpuratus* and *S. rosetta* caveolins lack a pin motif.

The three caveolin's N-terminal structured regions show drastic differences. Unlike the human Cav1 N-terminal region where the pin motif interacts with other protomers along the rim region and then is directly followed by the hook structure, the *S. purpuratus* caveolin N-terminal region extends outwards from the middle of the hydrophilic α-helical face until it reaches the hook

structure (a.a. 35-55) where it makes a ~180˚ turn (**Fig 8**). The N-terminus of the *S. rosetta* caveolin protomer begins with a short α-helix (a.a. 79-88) that starts about halfway up the C-terminal β-barrel and extends parallel to the disc on the presumed cytoplasmic side of the complex (**Fig 7G, H**). The N-terminal region continues to snake down the side of the C-terminal β-barrel (a.a. 216-230), makes a ~150˚ turn (a.a. 89-106) to continue along the face of the complex until it reaches the rim region. The *S. rosetta* caveolin protein, which like the *S. purpuratus* caveolin does not have a pin motif, does have a hook structure (a.a. 107-125) that makes a ~180˚ turn to facilitate extensive interactions between protomers (**Fig 8, Fig S10**). The human Cav1, *S. purpuratus* caveolin, and *S. rosetta* caveolin complexes all form extensive contacts among protomers. For the human Cav1, *S. purpuratus* caveolin complexes each protomer contacts two protomers to the left and two to the right (**Fig S10, A-K**). In the *S. rosetta* caveolin complex, additional contacts are made with the *i* + 5 and *i* - 5 protomers as the variable N-terminal region of *i* rises up the central β-barrel and contacts the C-terminal β strands of *i* + 5 and *i* – 5 (**Fig S10L**).

S. purpuratus and S. rosetta caveolin complexes form amphipathic discs that display increased curvature compared to the human Cav1 complex

A key feature of human Cav1 8S complex is the amphipathic nature of the disc (*27*) (**Fig 9A, D**). The *S. purpuratus* and *S. rosetta* caveolin complexes likewise contain distinct hydrophobic and hydrophilic faces (**Fig 9B, C, E, F**). As for the case of human Cav1, the hydrophobic face for both *S. purpuratus* and *S. rosetta* caveolin complexes, presumably representing the membrane-facing surfaces, are coated by detergent (**Fig 6I, 7I**). However, in contrast to the ring of glutamic acid residues present on the hydrophobic face of the human Cav1 complex, the putative membrane facing surface of *S. purpuratus* and *S. rosetta* caveolins lack any charged residues (**Fig 9B, C**). Predicted to face the cytosol, the hydrophilic faces of the complexes have an array of charged, but not conserved, residues across the caveolins (**Fig 9A-C**). Finally, the interior of the conserved central β-barrel is hydrophobic in the human Cav1, *S. purpuratus,* and *S. rosetta* caveolin complexes. While the human Cav1 β-barrel is capped by an inwardly facing charged Lys176, neither the *S. purpuratus* nor the *S. rosetta* caveolin β-barrels have this charged "cap" residue (**Fig 9D-F**).

The membrane facing surface of the human Cav1 8S complex is remarkably flat (*27*) (**Fig 9B**). In contrast, both the *S. purpuratus* and *S. rosetta* caveolin complexes are concave, with curvatures of ~17˚ and ~11˚, respectively (**Fig 9E, H**). Notably, this curvature is in the opposite direction one would expect for a protein that associates with the curved membrane invaginations seen in caveolae. We also found that 2D averages of vitrified *S. purpuratus* caveolin complexes exhibited variations of the degree of curvature, including examples of complexes with either concave or convex curvatures (**Fig 10A-C**). While we were unable to determine high-resolution structures from these averages, 3D variation analysis (3DVA) of the *S. purpuratus* caveolin complex results in components that capture a range of continuous complex conformations (**Movie 1**). Calculated using a filter resolution of 8Å, we show low-resolution structures representing the negative and positive values along the reaction coordinate for these components with the largest variance (**Fig 10E**). The 3DVA analysis show that the disc of the *S. purpuratus* caveolin complex can adopt a concave curvature on the membrane binding surface similar to what is seen in the 3.1 Å resolution map; however, the hydrophobic surface of the disc can also be flat. To accommodate these differences in curvature of the disc, the α-helices making up the spoke region rise in pitch towards the center of the complex, which also causes the β-barrel to be pushed "outwards" towards what would be the cytoplasm by ~4-5 Å (**Movie 1**). The variation of curvature that we see in the 2D averages of the *S. purpuratus* caveolin complex were not present in the 2D averages from the *S. rosetta* caveolin or human Cav1 samples (*27*). Further, 3DVA analysis failed to identify components displaying significant variations of curvature in either the *S. rosetta*

caveolin or human Cav1 discs. From these analyses, we conclude that the *S. purpuratus* caveolin complex is more flexible than either the human Cav1 or *S. rosetta* caveolin complexes under these experimental conditions.

Discussion

Caveolin homologues are present in a wide variety of organisms throughout Metazoa as well as in choanoflagellates, but little is known about the structural features of caveolins beyond human Cav1. Using a combination of computational and structural approaches, we now show that caveolins across species share the ability to assemble into disc-shaped multimers composed of spiraling α-helices and a central protruding parallel β-barrel. The discs have hydrophobic and hydrophilic faces that likely represent the membrane binding and cytoplasm orientated sides of the complex, respectively. Somewhat surprisingly considering the ability of AF2 to build disc-shaped complexes with different numbers of protomers, all three of the experimentally determined cryo-EM structures have 11-fold symmetry. While our analyses show many conserved features of caveolin oligomers, there are also some striking differences including the *S. purpuratus* and *S. rosetta* caveolin complexes having greater curvatures of the membrane-facing surface of the disc and smaller complex diameters. Additionally, the conformation of the N-terminal region of both *S. purpuratus* and *S. rosetta* caveolin complexes was distinct from the human Cav1 N-terminus. Finally, the *S. rosetta* caveolin complex's central β-barrel was 10 Å taller than those of either the human Cav1 or *S. purpuratus* caveolin complex. We conclude that caveolins across evolution adopt a conserved structural framework built around a disc-shaped cap composed of spiraling α-helices and a central protruding β-barrel, yet this common structure is adaptable enough to accommodate significant molecular variations. The conserved and variable regions of the caveolin structures could help explain why they have been maintained across evolution but also have many proposed functions in cells.

Classically, caveolins have been depicted as consisting of several major domains including a signature motif, scaffolding domain, oligomerization domain, intramembrane domain, and C-terminal region. However, based on our structure-based analyses of caveolins, we propose a new domain nomenclature. The computational and experimental structural analysis of caveolins shows these proteins can contain seven distinct structural units. Moving from the N- to C-terminus, these structural regions consist of a N-terminal variable region, pin motif, hook structure, scaffolding domain, spoke region, β-strand, and C-terminal variable region. Of these elements, the hook structure, scaffolding domain, and spoke region appear to be the most highly conserved structural features across most caveolins found in almost every clade that was examined. Our structures and the predicted models show that these regions of the protein contribute to oligomerization as well as help define the hydrophobic membrane facing surfaces of the complex. Interestingly, although these caveolin sequences have diverged evolutionarily for tens or even hundreds of millions of years, they all possess a remarkably similar signature motif. This attests to the functional significance of the caveolin signature motif, embedded in the region we now identify as the hook structure. The central parallel β-barrel is another reoccurring structural element. While similar in diameter and the hydrophobicity of central channel in the three caveolin cryo-EM structures, its length can vary. The conserved nature of the β-barrel raises questions about its biological functions. In human 8S Cav1 complex, the β-barrel contributes to the proper packing of protomers into regular disc-shaped complexes (*44*), and is disrupted by several disease-associated mutations (*27, 45-51*). The structures of *S. purpuratus* and *S. rosetta* caveolin also suggest similar roles in the proper packing of the complex. However, we also identified caveolins that are predicted to not form β-strands and thus would not have a β-barrel, including the *A. queenslandica* caveolin studied here. The absence of this structural motif could potentially help explain why the complexes *A. queenslandica* caveolin forms are less stable and homogenous disc than those formed by caveolins capable of assembling central β-barrels.

Whether the β-barrels fulfill additional physiological roles beyond complex organization and structural stability remains to be determined.

Other domains of caveolins are less well conserved across clades. For example, although the pin motif appears to be critical and essential for Type II-CAVs such as human Cav1 (*27*), it is probably the latest structural unit formed during the evolutionary process and is not essential for caveolins from other evolutionary branches to pack into oligomeric complexes. The length and composition of the N- and C-terminal unstructured regions are also highly variable across caveolins. While residues 1-48 of mammalian Cav1 are dispensable for caveolae formation (*3*), they fulfill several regulatory roles (*52-60*), contribute to the structural stability of the complex as analyzed by negative stain EM, and could potentially participate in curvature generation and/or condensate formation (*61, 62*). Given the strong divergence of the lengths and sequence identity of this region across other caveolins, this may represent the region of caveolins highly tuned for specific functions. Much less is known about the C-terminal variable region. Until now, this region has not been appreciated because in mammalian caveolins the C-terminal variable region is essentially absent (consisting of a single residue in human Cav1). It is also a very small stretch of amino acids in S. *purpuratus* and *S. rosetta* caveolin (a.a. 153-159 and 232-233, respectively) which are missing in the cryo-EM density maps for these caveolins. However, sequence analysis indicates that in other caveolins, this region may be upwards of 45 residues in length. Because this region is expected to project into the cytoplasm, we speculate that it could represent a binding site for cytosolic proteins.

The finding that caveolin-related proteins are not limited to Metazoa but are also found in choanoflagellates implies that the proteins evolved independently in the two lineages for the past billion years (*5*). Despite this, our structural analysis shows that the choanoflagellate caveolin from *S. rosetta* shares many structural features with human Cav1, including the ability to form amphipathic disc-like oligomers with a central parallel β-barrel. This suggests that a membrane-embedded disc-shaped oligomer may represent one of the most ancient functional units of the protein. The *S. rosetta* caveolin protein also contains an elaborate, more extensively structured N-terminal region that forms an additional helical region not observed in the metazoan structures, and the curvature of its membrane facing surface is markedly different from that of human Cav1. Furthermore, choanoflagellates, like non-vertebrate metazoans, do not appear to have cavins in their genomes and therefore their caveolins would not be expected to assemble into caveolae similar to those observed in vertebrates. These findings raise the interesting question of what functional roles caveolins fulfill in choanoflagellates and how these relate to their functions in mammals. A potential clue is that several binding partners and signaling pathways that caveolins have been linked to in mammals are also found in choanoflagellates. Specific examples include cell adhesion proteins and Src (*31, 63, 64*). Another possibility is that they buffer changes in membrane tension by a cavin-independent mechanism (*65*). As experimental approaches to study the cell biology of choanoflagellates continue to advance (*66*), it should be possible to test the structure-function relationship of this evolutionarily diverse forms of the protein in the future.

There are several limitations to our findings. First, AF2, while an exceptional tool, has some weaknesses. As one example, while AF2 is able to predict the basic secondary structure and the overall organization of caveolins into spirally packed discs, it does less well in predicting the structure of the N- and C-terminal domains of *S. purpuratus* and *S. rosetta* caveolins. AF2.2 also appears to "overpredict" the assembly of certain caveolins, such as human Cav2. It is well-known that human Cav2 cannot be incorporated into a stable 8S complex unless Cav1 is present (*67-70*), yet AF2.2 still predicts a homotypic 11-mer with regular shapes (as does AF3) (**Supplementary File 2**). Interestingly, AF2.1 could predict the "deficiency" of Cav2 in assembly and struggled to assemble Cav2 complexes as a closed structure (**Supplementary File 1**). We anticipate that as more experimentally derived caveolin structures continue to be added to the AF

training database, future AF predictions will become more accurate in this regard. How the oligomeric state of caveolin complexes is controlled is also uncertain. We found that human Cav1, *S. purpuratus* caveolin, and *S. rosetta* caveolin all form 11-mer complexes; however, the oligomerization state of the complexes may be influenced by expression and purification conditions of our experiments. Indeed, there is experimental evidence that suggests varied oligomeric states of mammalian caveolins may assemble in cells (*37, 71-74*). Consistent with this, AF2 predicts additional oligomeric states can exist in addition to the experimentally observed 11-mers, as illustrated in our previous work (*75*) and the current study. However, some of the oligomeric states predicted by AF2 appear less likely to exist due to energetic strains (*75*). One interesting possibility is that membrane lipid composition plays a role controlling the oligomerization of caveolin complexes (*76*), a hypothesis that remains to be fully tested. Finally, in contrast to human Cav1 complex's flat membrane-facing surface, both the *S. purpuratus* and *S. rosetta* caveolin complexes exhibited concave curvature. Through 2D classification and 3DVA, we found that the *S. purpuratus* caveolin complex's membrane-facing surface samples additional curvatures as well. The differences in curvature and the ability of *S. purpuratus* caveolin to adopt multiple curvature conformations raises the interesting possibility that caveolins may be optimized to bend membranes into different shapes and/or undergo conformational changes in response to changes in membrane tension.

In summary, we conclude that that the ability of caveolins to assemble into amphipathic discs represents an ancient, unique mode of protein-lipid membrane interactions that predates the emergence of metazoans. Given these new insights, it should now become possible to uncover how caveolins control cellular function at a molecular level across an evolutionary scale.

Materials and Methods

AlphaFold2 Predictions

Predicted structures of caveolin homo-oligomers were generated using AF2.1 or AF2.2 by systematically increasing the number of monomer input sequences until the upper limit of the residues that could be analyzed was reached. AlphaFold v2.1.0 predictions were performed using a Colab notebook named "AlphaFold2_advanced (AlphaFold2_advanced.ipynb - Colaboratory (google.com)" with default settings. AlphaFold v2.2.0 predictions were performed using default settings via another Colab notebook named "alphafold21_predict_colab" provided by ChimeraX daily builds version (ChimeraX 1.4.0). Version v2.2.0 includes updated AlphaFold-Multimer model parameters. See https://github.com/deepmind/alphafold/releases for a description of other differences in AF2.1.0 versus AF2.2.0. Due to the upper limit in the number of residues that could be analyzed by AF2.1, where indicated, caveolin sequences were truncated to exclude the predicted N-terminal disordered regions. Unless otherwise stated, the rank model 1 of the 5 models output for each prediction is shown. Confidence levels of the predictions are rendered on the models using pLDDT (predicted local-distance difference test) values (*32*).

Conservation analysis

Conservation analysis was based on protein sequences alignments collected from UniProt of 72 metazoan caveolin sequences from 13 representative species of different phylum/superphylum via EMBL-EBI online tool MAFFT (Multiple Alignment using Fast Fourier Transform) (MAFFT < Multiple Sequence Alignment < EMBL-EBI) and displayed on 3D structure by using Chimera v1.15. The figures were analyzed, rendered, and exported with Chimera v1.15, ChimeraX 1.4.0 (daily builds version), or ChimeraX1.3.

Sequence alignments

Clustal Omega was used to perform sequence alignments. Jalview 2.11.2.4 (*77*) was used for alignment image typesetting and exporting.

HMMER searches

For HMMER searches, a hidden Markov model (HMM) profile was built directly from a MAFFT alignment of previously reported truncated caveolin sequences (*3*) using hmmbuild from the HMMER package v 3.1b2 (*78*). The profile was then searched against the *S. rosetta* ATCC 50818 NCBI reference genome and *A. queenslandica* genome v1.1 from EnsemblMetazoa using hmmsearch. To identify putative ctenophore caveolins we used blastp v2.10.0+ (*79*) with the human protein sequences as queries, and HMMer with the above models to query the Hormiphora californensis (*42*) and Bolinopsis microptera (*41*) genomes (GCA_020137815.1 and GCF_026151205.1) and transcriptomes (TSA GHXS00000000 and GCF_026151205.1 annotation).

Phylogenetic analyses

The selected caveolin homologues were aligned with MAFFT v7.310 (*80, 81*). The alignment was then truncated to a region corresponding to the residues 54–158 in human CAV1 using a simple Python script (https://doi.org/10.5281/zenodo.6562402). Gaps were removed before combining the sequences with those from the supplementary information of Kirkham et al (2008) (*3*) and re-aligning with MAFFT. ProtTest3 (version 3.4.2) (*82*) was then used to determine the best model of evolution (LG+I+Γ). Finally, a maximum likelihood phylogeny was inferred using RAxML (version 8.2.11; (*83, 84*) with 1000 rapid bootstraps. The trimmed tree was produced in RStudio using the *ape* package (*85*). Trees were rendered using FigTree v1.4.4 and prepared for publication using Inkscape (version 1.0.1).

Expression and purification of recombinant caveolins in *E. coli*

Protein expression and purification were performed following previously described protocols with minor modifications (*44*). Genes encoding the entire caveolin polypeptide of *A. queenslandica* (UniProt Accession A0A1X7UHP5), *S. purpuratus* (UniProt Accession A0A7M7T4C2) and *S. Rosetta* (UniProt Accession F2U793) were synthesized by Genscript, NJ and subcloned into the pET20b(+) vector (Novagen) using NdeI and XhoI (New England Biolabs, MA) restriction sites. The primers used (Integrated DNA Technologies, IA) were as follows: *A. queenslandica* forward gcggccCATATGCCTCCACCCCCTCCCCCG and reverse gcggccCTCGAGtcttttgaagataatggccacag; *S. purpuratus* forward gcggccCATATGGAACTGATCCATCCTG and reverse gcggccCTCGAGcacctggctggtcttaacatcagagc and *S. Rosetta* forward gcggccCATATGAGCTACCAC and reverse gcggccCTCGAGgtccttcagttccttgtgg. Resulting plasmids were verified by Sanger sequencing (Genewiz/Azenta Life Sciences, NJ). Caveolin proteins were expressed in *E. coli* BL21 using the auto-induction expression system (*44*). Initially, an MDG starter culture of bacteria was incubated at 37°C and 250 rpm for 20 hours. Subsequently, the culture was expanded using auto-inducing ZYM-5052 media at 25°C and 300 rpm for 24 hours. The *E. coli* cells were then washed with 0.9% NaCl and resuspended in a buffer containing 200 mM NaCl and 20 mM Tris–HCl, pH 8.0. Bacterial cells were homogenized using a French press pressure homogenizer, with 1 mM PMSF and DTT added immediately before homogenization. To remove large cell debris, the homogenate was centrifuged at 9000 rpm for 15 minutes at 4°C. Total membranes were subsequently pelleted by centrifugation at 40,000 rpm (Ti-45 rotor, Beckman Coulter) for 1 hour at 4°C. The membrane pellets were then homogenized using a Dounce tissue grinder in a buffer consisting of 200 mM NaCl, 20 mM Tris–HCl (pH 8.0), and 1 mM DTT. To solubilize caveolin proteins from the membranes, a 10% n-Dodecyl- β-D-Maltopyranoside (C12M) (Anatrace) stock solution was added to the membrane homogenate to

a final concentration of 2%, and the mixture was gently stirred for 2 hours at 4°C. Insoluble material was removed by centrifugation at 42,000 rpm (Ti-50.2 rotor) for 35 minutes, and the supernatant was subjected to Nickel Sepharose affinity purification. After washing the resin using 8 column volumes of a buffer composed of 200 mM NaCl, 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 0.05% C12M, and 60 mM imidazole, the protein was eluted by increasing the imidazole concentration to 300 mM. The eluate containing caveolin was concentrated and further purified by size-exclusion chromatography using a Superose 6 Increase 10/300 GL column (GE Healthcare) in a buffer containing 200 mM NaCl, 20 mM Tris–HCl (pH 8.0), 1 mM DTT, and 0.05% C12 M.

Negative stain EM and data processing

Caveolin samples were prepared for negative stain EM using established methods (*86*). Briefly, a PELCO easiGlow™ glow discharge unit (Fresno, CA, USA) was used to glow discharge 200-mesh copper grids covered with carbon-coated collodion film (EMS, Hatfield, PA, USA) for 30 s at 10 mA. Evolutionary caveolin samples (3.5µL) were adsorbed to the grids and incubated for 1 minute at room temperature. Samples were first washed with two drops of water and then stained with two drops of 0.7% (w/v) uranyl formate (EMS, Hatfield, PA, USA). The samples were then blotted until dry. A Morgagni transmission electron microscope operated at an accelerating voltage of 100 kV (Thermo Fisher Scientific, Waltham, MA, USA) was used to image the samples at a nominal magnification of 22,000x (2.1 Å per pixel).

Negative stain EM datasets were collected with a Tecnai Spirit T12 transmission electron microscope operated at 120 kV (Thermo Fisher Scientific, Waltham, MA, USA). Images were collected at a nominal magnification of 30,000x (2.086 Å per pixel). Data were collected using SerialEM v4.0.8 software (*87*) on a 4 k x 4 k Rio camera (Gatan, Pleasanton, CA) with a -2.2 µm defocus value. All image processing was performed in Relion 4.0.0 (*88*). Approximately 1,000 particles were manually selected and 2D classified, and clear resulting classes were selected and used as references for particle picking on all micrographs. Particles were extracted with a 144-pixel box size (30 nm x 30 nm boxes). The extracted particles were 2D classified into 20 classes. The human Cav1, *S. purpuratus* caveolin, *A. queenslandica* caveolin, and *S. rosetta* caveolin datasets had 39,786, 124,021, 66,856, and 94,597 particles respectively.

Cryo-EM sample preparation

For single particle cryo-EM, 4 µL of the protein sample (*S. purpuratus* caveolin: ~0.02 mg/mL; *S. rosetta* caveolin: ~0.04 mg/mL) was applied to a Quantifoil R 2/2, 200-mesh Cu grid with an ultrathin carbon layer (Electron Microscopy Services) that was glow-discharged for 30 s at 5 mA. Following a 30 s incubation, the sample was removed, and another 4 µL of protein sample was applied to the same grid. After a second 30 s incubation, the sample was blotted for 5 s with a blot force of 10 and was then plunge frozen in a slurry of liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). The chamber was kept at 4ºC with 100% humidity.

Cryo-EM data collection

Micrographs of the *S. purpuratus* caveolin sample were collected on a Titan Krios transmission electron microscope (Thermo Fisher Scientific) operated at 300 kV and equipped with a K3 direct detection camera with a BioQuantum energy filter used with a slit width of 20 eV (Gatan). Images were collected at a nominal magnification of 81,000 x (1.11 Å/ pixel). Two datasets were collected of the *S. purpuratus* caveolin sample with 10,127 and 12,803 micrographs using SerialEM v4.0.3 software with a total dose of 60 e-/Å2 and a defocus range of -0.5 µm to -3 µm (*87*).

Images of the *S. rosetta* caveolin sample were collected on a Titan Krios transmission electron microscope (Thermo Fisher Scientific). The microscope was operated at 300 kV and equipped with a K3 direct detection camera with a BioQuantum energy filter used with a slit width of 20 eV (Gatan). Images were collected at a nominal magnification of 105,000 x (0.87Å/ pixel). A total of 18,417 micrographs were collected using SerialEM v4.0.3 software with a total dose of 59.17 e-/Å2 and a defocus range of -0.5 µm to -3 µm (*87*).

Image processing of *S. purpuratus* caveolin complex

All image processing, 2D classification, and 3D refinements were performed in cryoSPARC v4.2.1 and v4.3.1 (*89*). Two independently collected datasets of 10,127 and 12,803 movies respectively (22,930 total movies) were corrected for local-beam induced drift using patch motion correction. Patch CTF estimation was used to estimate the local CTF parameters. Following exposure curation to keep only micrographs with a CTF estimation of ≤5Å, 22,845 total micrographs were subjected to circular blob picking. After blob picking, 13,044,185 initial picks were extracted in 352-pixel^2 boxes (390.72 Å x 390.72 Å). Iterative 2D classification resulted in 200,411 particles. These particles were input into a two-class *ab initio* 3D reconstruction. The 135,462 particles that contributed to the reconstruction yielding a better-defined caveolin complex were subjected to a further round of 2D classification, resulting in 80,245 particles. These particles were used for another two-class *ab initio* 3D reconstruction. The 44,388 particles contributing to the better resolved *ab initio* reconstruction were input into a single class *ab initio* 3D reconstruction with higher initial and maximum resolution parameters. The resulting *ab initio* 3D reconstruction and the 44,388 particles contributing to it were used as inputs for C1 and C11 non-uniform refinements. The C1 non-uniform refinement resulted in a map with a resolution of 6.6Å, and the C11 non-uniform refinement produced a map with a resolution of 3.2Å. Local masks were generated in ChimeraX and Relion 4.0 to mask out the detergent micelle (*88, 90*). These masks were then used to run local refinements with C1 and C11 symmetry. The C1 refinement resulted in a map of 6.2 Å resolution, and the C11 refinement reached a final resolution of 3.1 Å. All 3DVA was carried out in cryoSPARC v4.3.1 (*91*).

Image processing of *S. rosetta* caveolin complex

18,417 movies were motion corrected and CTF estimated using Warp (*92*). Particles were picked using the BoxNet2 centralized neural network (CNN) in Warp, and 1,488,129 initial particles were extracted with a box size of 300-pixel^2 (261Å x 261Å). Iterative 2D classification of particles in cryoSPARC v4.2.1 resulted in 66,202 particles (*89*). These particles were then input to a single-class *ab initio* 3D reconstruction that resulted in a reconstruction exhibiting 11 spiraling α-helices. The 66,202 particles contributing to the *ab initio* reconstruction were used for a non-uniform refinement with C1 symmetry that resulted in a map with a resolution of 3.8 Å. Imposing C11 symmetry for non-uniform refinement of the same particles resulted in a 3.4 Å map. Using ChimeraX and Relion 4.0, masks were generated to mask out the detergent micelle (*88, 90*). Local refinement using these masks resulted in a C1 refinement with a final resolution of 3.0 Å and a C11 refinement with a final resolution of 2.9 Å.

Model building, refinement, and validation

Models of the *S. purpuratus* and *S. rosetta* caveolin proteins were built using ModelAngelo with the respective sequences input into the program (*93*). The output models resulting from ModelAngelo were then further refined using ISOLDE within ChimeraX and Phenix real-space refinement (*90, 94, 95*). The final model of *S. purpuratus* caveolin included residues 29-152, and the model of *S. rosetta* caveolin contained residues 79-231. Both models were validated within Phenix (**Table S2**) (*95*). Maps and models are deposited in the EMDB and PDB (*S. purpuratus*

caveolin complex*: EMD-XXXXX, PDB-XXXXX; S. rosetta* caveolin complex*: EMD-XXXXX, PDB-XXXXX).*

References

1. M. M. Attwood, A. Krishnan, M. S. Almen, H. B. Schioth, Highly diversified expansions shaped the evolution of membrane bound proteins in metazoans. *Sci Rep* **7**, 12387 (2017).

2. M. C. Field, C. Gabernet-Castello, J. B. Dacks, Reconstructing the evolution of the endocytic system: insights from genomics and molecular cell biology. *Adv Exp Med Biol* **607**, 84-96 (2007).

3. M. Kirkham, S. J. Nixon, M. T. Howes, L. Abi-Rached, D. E. Wakeham, M. Hanzal-Bayer, C. Ferguson, M. M. Hill, M. Fernandez-Rojo, D. A. Brown, J. F. Hancock, F. M. Brodsky, R. G. Parton, Evolutionary analysis and molecular dissection of caveola biogenesis. *J Cell Sci* **121**, 2075-2086 (2008).

4. E. Spisni, V. Tomasi, A. Cestaro, S. C. Tosatto, Structural insights into the function of human caveolin 1. *Biochem Biophys Res Commun* **338**, 1383-1390 (2005).

5. J. F. H. Strassert, I. Irisarri, T. A. Williams, F. Burki, A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat Commun* **12**, 1879 (2021).

6. T. M. Williams, M. P. Lisanti, The Caveolin genes: from cell biology to medicine. *Ann Med* **36**, 584-595 (2004).

7. M. Bastiani, R. G. Parton, Caveolae at a glance. *J Cell Sci* **123**, 3831-3836 (2010).

8. R. G. Parton, Caveolae: structure, function, and relationship to disease. *Annu Rev Cell Dev Biol* **34**, 111-136 (2018).

9. J. H. Chidlow, Jr., W. C. Sessa, Caveolae, caveolins, and cavins: complex control of cellular signalling and inflammation. *Cardiovasc Res* **86**, 219-225 (2010).

10. M. A. Del Pozo, F. N. Lolo, A. Echarri, Caveolae: Mechanosensing and mechanotransduction devices linking membrane trafficking to mechanoadaptation. *Curr Opin Cell Biol* **68**, 113-123 (2020).

11. R. G. Parton, M. A. del Pozo, Caveolae as plasma membrane sensors, protectors and organizers. *Nat Rev Mol Cell Biol* **14**, 98-112 (2013).

12. P. Nassoy, C. Lamaze, Stressing caveolae new role in cell mechanics. *Trends Cell Biol* **22**, 381-389 (2012).

13. C. Lamaze, N. Tardif, M. Dewulf, S. Vassilopoulos, C. M. Blouin, The caveolae dress code: structure and signaling. *Curr Opin Cell Biol* **47**, 117-125 (2017).

14. J. B. Dacks, M. S. Robinson, Outerwear through the ages: evolutionary cell biology of vesicle coats. *Curr Opin Cell Biol* **47**, 108-116 (2017).

15. B. Razani, S. E. Woodman, M. P. Lisanti, Caveolae: from cell biology to animal physiology. *Pharmacol Rev* **54**, 431-467 (2002).

16. R. G. Parton, V. Tillu, K. A. McMahon, B. M. Collins, Key phases in the formation of caveolae. *Curr Opin Cell Biol* **71**, 7-14 (2021).

17. R. G. Parton, K. A. McMahon, Y. Wu, Caveolae: formation, dynamics, and function. *Curr Opin Cell Biol* **65**, 8-16 (2020).

18. A. Pol, F. Morales-Paytuvi, M. Bosch, R. G. Parton, Non-caveolar caveolins - duties outside the caves. *J Cell Sci* **133**, (2020).

19. O. Kovtun, V. A. Tillu, N. Ariotti, R. G. Parton, B. M. Collins, Cavin family proteins and the assembly of caveolae. *J Cell Sci* **128**, 1269-1278 (2015).

20. N. Ariotti, R. G. Parton, SnapShot: caveolae, caveolins, and cavins. *Cell* **154**, 704-704 e701 (2013).

21. B. P. Head, P. A. Insel, Do caveolins regulate cells by actions outside of caveolae? *Trends Cell Biol* **17**, 51-57 (2007).

22. C. G. Hansen, B. J. Nichols, Exploring the caves: cavins, caveolins and caveolae. *Trends Cell Biol* **20**, 177-186 (2010).

23. Z. Tang, T. Okamoto, P. Boontrakulpoontawee, T. Katada, A. J. Otsuka, M. P. Lisanti, Identification, sequence, and expression of an invertebrate caveolin gene family from the nematode Caenorhabditis elegans. Implications for the molecular evolution of mammalian caveolin genes. *J Biol Chem* **272**, 2437-2445 (1997).

24. P. G. Frank, M. P. Lisanti, Zebrafish as a novel model system to study the function of caveolae and caveolin-1 in organismal biology. *Am J Pathol* **169**, 1910-1912 (2006).

25. P. Bhattachan, J. Rae, H. Yu, W. Jung, J. Wei, R. G. Parton, B. Dong, Ascidian caveolin induces membrane curvature and protects tissue integrity and morphology during embryogenesis. *FASEB J* **34**, 1345-1361 (2020).

26. B. Razani, D. S. Park, Y. Miyanaga, A. Ghatpande, J. Cohen, X. B. Wang, P. E. Scherer, T. Evans, M. P. Lisanti, Molecular cloning and developmental expression of the caveolin gene family in the amphibian Xenopus laevis. *Biochemistry* **41**, 7914-7924 (2002).

27. J. C. Porta, B. Han, A. Gulsevin, J. M. Chung, Y. Peskova, S. Connolly, H. S. Mchaourab, J. Meiler, E. Karakas, A. K. Kenworthy, M. D. Ohi, Molecular architecture of the human caveolin-1 complex. *Sci Adv* **8**, eabn7232 (2022).

28. M. D. Ohi, A. K. Kenworthy, Emerging insights into the molecular architecture of caveolin-1. *J Membr Biol* **255**, 375–383 (2022).

29. M. Doktorova, S. Daum, J. Ebenhan, S. Neudorf, B. Han, S. Sharma, P. Kasson, K. Levental, K. Bacia, A. K. Kenworthy, I. Levental, Caveolin assemblies displace one bilayer leaflet to organize and bend membranes. *bioRxiv*, 2024.2008.2028.610209 (2024).

30. A. Echarri, M. A. Del Pozo, Caveolae. *Curr Biol* **22**, R114-116 (2012).

31. K. Shalchian-Tabrizi, M. A. Minge, M. Espelund, R. Orr, T. Ruden, K. S. Jakobsen, T. Cavalier-Smith, Multigene phylogeny of choanozoa and the origin of animals. *PLoS One* **3**, e2098 (2008).

32. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

33. A. Gulsevin, B. Han, J. C. Porta, H. S. Mchaourab, J. Meiler, A. K. Kenworthy, Template-free prediction of a new monotopic membrane protein fold and oligomeric assembly by Alphafold2. *Biophys J* **in press**, (2022).

34. S. Li, T. Okamoto, M. Chun, M. Sargiacomo, J. E. Casanova, S. H. Hansen, I. Nishimoto, M. P. Lisanti, Evidence for a regulated interaction between heterotrimeric G proteins and caveolin. *J Biol Chem* **270**, 15693-15701 (1995).

35. A. Schlegel, R. B. Schwab, P. E. Scherer, M. P. Lisanti, A role for the caveolin scaffolding domain in mediating the membrane attachment of caveolin-1. The caveolin scaffolding domain is both necessary and sufficient for membrane binding in vitro. *J Biol Chem* **274**, 22660-22667 (1999).

36. K. T. Root, S. M. Plucinsky, K. J. Glover, Recent progress in the topology, structure, and oligomerization of caveolin: a building block of caveolae. *Curr Top Membr* **75**, 305-336 (2015).

37. M. Sargiacomo, P. E. Scherer, Z. Tang, E. Kubler, K. S. Song, M. C. Sanders, M. P. Lisanti, Oligomeric structure of caveolin: implications for caveolae membrane organization. *Proc Natl Acad Sci U S A* **92**, 9407-9411 (1995).

38. R. G. Parton, M. Hanzal-Bayer, J. F. Hancock, Biogenesis of caveolae: a structural model for caveolin-induced domain formation. *J Cell Sci* **119**, 787-796 (2006).

39. S. R. Fairclough, Z. Chen, E. Kramer, Q. Zeng, S. Young, H. M. Robertson, E. Begovic, D. J. Richter, C. Russ, M. J. Westbrook, G. Manning, B. F. Lang, B. Haas, C. Nusbaum, N. King, Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. *Genome Biol* **14**, R15 (2013).

40. M. Srivastava, O. Simakov, J. Chapman, B. Fahey, M. E. Gauthier, T. Mitros, G. S. Richards, C. Conaco, M. Dacre, U. Hellsten, C. Larroux, N. H. Putnam, M. Stanke, M. Adamska, A. Darling, S. M. Degnan, T. H. Oakley, D. C. Plachetzki, Y. Zhai, M. Adamski, A. Calcino, S. F. Cummins, D. M. Goodstein, C. Harris, D. J. Jackson, S. P. Leys, S. Shu, B. J. Woodcroft, M. Vervoort, K. S. Kosik, G. Manning, B. M. Degnan, D. S. Rokhsar, The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature* **466**, 720-726 (2010).

41. D. T. Schultz, S. H. D. Haddock, J. V. Bredeson, R. E. Green, O. Simakov, D. S. Rokhsar, Ancient gene linkages support ctenophores as sister to other animals. *Nature* **618**, 110-117 (2023).

42. D. T. Schultz, W. R. Francis, J. D. McBroome, L. M. Christianson, S. H. D. Haddock, R. E. Green, A chromosome-scale genome assembly and karyotype of the ctenophore Hormiphora californensis. *G3 (Bethesda)* **11**, (2021).

43. T. C. Terwilliger, D. Liebschner, T. I. Croll, C. J. Williams, A. J. McCoy, B. K. Poon, P. V. Afonine, R. D. Oeffner, J. S. Richardson, R. J. Read, P. D. Adams, AlphaFold predictions are valuable hypotheses, and accelerate but do not replace experimental structure determination. *bioRxiv*, 2022.2011.2021.517405 (2023).

44. B. Han, J. C. Porta, J. L. Hanks, Y. Peskova, E. Binshtein, K. Dryden, D. P. Claxton, H. S. Mchaourab, E. Karakas, M. D. Ohi, A. K. Kenworthy, Structure and assembly of CAV1 8S complexes revealed by single particle electron microscopy. *Sci Adv* **6**, eabc6185 (2020).

45. E. D. Austin, L. Ma, C. LeDuc, E. Berman Rosenzweig, A. Borczuk, J. A. Phillips, 3rd, T. Palomero, P. Sumazin, H. R. Kim, M. H. Talati, J. West, J. E. Loyd, W. K. Chung, Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. *Circ Cardiovasc Genet* **5**, 336-343 (2012).

46. A. Garg, M. Kircher, M. Del Campo, R. S. Amato, A. K. Agarwal, Whole exome sequencing identifies de novo heterozygous CAV1 mutations associated with a novel neonatal onset lipodystrophy syndrome. *Am J Med Genet A* **167**, 1796-1806 (2015).

47. I. Schrauwen, S. Szelinger, A. L. Siniard, A. Kurdoglu, J. J. Corneveaux, I. Malenica, R. Richholt, G. Van Camp, M. De Both, S. Swaminathan, M. Turk, K. Ramsey, D. W. Craig, V. Narayanan, M. J. Huentelman, A frame-shift mutation in CAV1 is associated with a severe neonatal progeroid and lipodystrophy syndrome. *PLoS One* **10**, e0131797 (2015).

48. B. Han, C. A. Copeland, Y. Kawano, E. B. Rosenzweig, E. D. Austin, L. Shahmirzadi, S. Tang, K. Raghunathan, W. K. Chung, A. K. Kenworthy, Characterization of a caveolin-1 mutation associated with both pulmonary arterial hypertension and congenital generalized lipodystrophy. *Traffic* **17**, 1297-1312 (2016).

49. C. A. Copeland, B. Han, A. Tiwari, E. D. Austin, J. E. Loyd, J. D. West, A. K. Kenworthy, A disease-associated frameshift mutation in caveolin-1 disrupts caveolae formation and function through introduction of a de novo ER retention signal. *Mol Biol Cell* **28**, 3095-3111 (2017).

50. G. Marsboom, Z. Chen, Y. Yuan, Y. Zhang, C. Tiruppathi, J. E. Loyd, E. D. Austin, R. F. Machado, R. D. Minshall, J. Rehman, A. B. Malik, Aberrant caveolin-1-mediated Smad signaling and proliferation identified by analysis of adenine 474 deletion mutation (c.474delA) in patient fibroblasts: a new perspective on the mechanism of pulmonary hypertension. *Mol Biol Cell* **28**, 1177-1185 (2017).

51. S. Plucinsky, K. Glover, The C-terminal domain of caveolin-1 and pulmonary arterial hypertension: An emerging relationship. *J Rare Dis Res Treat.* **2**, 44-48 (2017).

52. P. K. Fang, K. R. Solomon, L. Zhuang, M. Qi, M. McKee, M. R. Freeman, P. C. Yelick, Caveolin-1alpha and -1beta perform nonredundant roles in early vertebrate development. *Am J Pathol* **169**, 2209-2222 (2006).

53. P. Kirchner, M. Bug, H. Meyer, Ubiquitination of the N-terminal region of caveolin-1 regulates endosomal sorting by the VCP/p97 AAA-ATPase. *J Biol Chem* **288**, 7363-7372 (2013).

54. J. R. Glenney, Jr., L. Zokas, Novel tyrosine kinase substrates from Rous sarcoma virus-transformed cells are present in the membrane skeleton. *J Cell Biol* **108**, 2401-2408 (1989).

55. S. Li, R. Seitz, M. P. Lisanti, Phosphorylation of caveolin by src tyrosine kinases. The alpha-isoform of caveolin is selectively phosphorylated by v-Src in vivo. *J. Biol. Chem.* **271**, 3863-3868 (1996).

56. J. P. Vainonen, N. Aboulaich, M. V. Turkina, P. Stralfors, A. V. Vener, N-terminal processing and modifications of caveolin-1 in caveolae from human adipocytes. *Biochem Biophys Res Commun* **320**, 480-486 (2004).

57. C. Corley Mastick, A. R. Sanguinetti, J. H. Knesek, G. S. Mastick, L. F. Newcomb, Caveolin-1 and a 29-kDa caveolin-associated protein are phosphorylated on tyrosine in cells expressing a temperature-sensitive v-Abl kinase. *Exp Cell Res* **266**, 142-154 (2001).

58. C. C. Mastick, A. R. Saltiel, Insulin-stimulated tyrosine phosphorylation of caveolin is specific for the differentiated adipocyte phenotype in 3T3-L1 cells. *J Biol Chem* **272**, 20706-20714 (1997).

59. R. D. Parr, G. G. Martin, H. A. Hostetler, M. E. Schroeder, K. D. Mir, A. B. Kier, J. M. Ball, F. Schroeder, A new N-terminal recognition domain in caveolin-1 interacts with sterol carrier protein-2 (SCP-2). *Biochemistry* **46**, 8301-8314 (2007).

60. M. A. Doucey, F. C. Bender, D. Hess, J. Hofsteenge, C. Bron, Caveolin-1 interacts with the chaperone complex TCP-1 and modulates its protein folding activity. *Cell Mol Life Sci* **63**, 939-948 (2006).

61. D. J. Busch, J. R. Houser, C. C. Hayden, M. B. Sherman, E. M. Lafer, J. C. Stachowiak, Intrinsically disordered proteins drive membrane curvature. *Nat Commun* **6**, 7875 (2015).

62. F. Yuan, H. Alimohamadi, B. Bakka, A. N. Trementozzi, K. J. Day, N. L. Fawzi, P. Rangamani, J. C. Stachowiak, Membrane bending by protein phase separation. *Proc Natl Acad Sci U S A* **118**, e2017435118 (2021).

63. N. King, C. T. Hittinger, S. B. Carroll, Evolution of key cell signaling and adhesion protein families predates animal origins. *Science* **301**, 361-363 (2003).

64. Y. Segawa, H. Suga, N. Iwabe, C. Oneyama, T. Akagi, T. Miyata, M. Okada, Functional development of Src tyrosine kinases during evolution from a unicellular ancestor to multicellular animals. *Proc Natl Acad Sci U S A* **103**, 12021-12026 (2006).

65. F. N. Lolo, N. Walani, E. Seemann, D. Zalvidea, D. M. Pavon, G. Cojoc, M. Zamai, C. Viaris de Lesegno, F. Martinez de Benito, M. Sanchez-Alvarez, J. J. Uriarte, A. Echarri, D. Jimenez-Carretero, J. C. Escolano, S. A. Sanchez, V. R. Caiolfa, D. Navajas, X. Trepat, J. Guck, C. Lamaze, P. Roca-Cusachs, M. M. Kessels, B. Qualmann, M. Arroyo, M. A. Del Pozo, Caveolin-1 dolines form a distinct and rapid caveolae-independent mechanoadaptation system. *Nat Cell Biol* **25**, 120-133 (2023).

66. D. S. Booth, H. Szmidt-Middleton, N. King, Transfection of choanoflagellates illuminates their cell biology and the ancestry of animal septins. *Mol Biol Cell* **29**, 3026-3038 (2018).

67. S. Li, F. Galbiati, D. Volonte, M. Sargiacomo, J. A. Engelman, K. Das, P. E. Scherer, M. P. Lisanti, Mutational analysis of caveolin-induced vesicle formation. Expression of caveolin-1 recruits caveolin-2 to caveolae membranes. *FEBS Lett* **434**, 127-134 (1998).

68. I. Parolini, M. Sargiacomo, F. Galbiati, G. Rizzo, F. Grignani, J. A. Engelman, T. Okamoto, T. Ikezu, P. E. Scherer, R. Mora, E. Rodriguez-Boulan, C. Peschle, M. P. Lisanti, Expression of caveolin-1 is required for the transport of caveolin-2 to the plasma membrane. Retention of caveolin-2 at the level of the golgi complex. *J Biol Chem* **274**, 25718-25725 (1999).

69. R. Mora, V. L. Bonilha, A. Marmorstein, P. E. Scherer, D. Brown, M. P. Lisanti, E. Rodriguez-Boulan, Caveolin-2 localizes to the Golgi complex but redistributes to plasma membrane, caveolae, and rafts when co-expressed with caveolin-1. *J. Biol. Chem.* **274**, 25708-25717 (1999).

70. A. Hayer, M. Stoeber, C. Bissig, A. Helenius, Biogenesis of caveolae: stepwise assembly of large caveolin and cavin complexes. *Traffic* **11**, 361-382 (2010).

71. I. Fernandez, Y. Ying, J. Albanesi, R. G. Anderson, Mechanism of caveolin filament assembly. *Proc Natl Acad Sci U S A* **99**, 11193-11198 (2002).

72. S. Monier, R. G. Parton, F. Vogel, J. Behlke, A. Henske, T. V. Kurzchalia, VIP21-caveolin, a membrane protein constituent of the caveolar coat, oligomerizes in vivo and in vitro. *Mol. Biol. Cell* **6**, 911-927 (1995).

73. I. M. Khater, F. Meng, T. H. Wong, I. R. Nabi, G. Hamarneh, Super resolution network analysis defines the molecular architecture of caveolae and caveolin-1 scaffolds. *Sci Rep* **8**, 9009 (2018).

74. I. M. Khater, Q. Liu, K. C. Chou, G. Hamarneh, I. R. Nabi, Super-resolution modularity analysis shows polyhedral caveolin-1 oligomers combine to form scaffolds and caveolae. *Sci Rep* **9**, 9888 (2019).

75. A. Gulsevin, B. Han, J. C. Porta, H. S. McHaourab, J. Meiler, A. K. Kenworthy, Template-free prediction of a new monotopic membrane protein fold and assembly by AlphaFold2. *Biophys J* **122**, 2041-2052 (2023).

76. A. K. Kenworthy, B. Han, N. Ariotti, R. G. Parton, The Role of Membrane Lipids in the Formation and Function of Caveolae. *Cold Spring Harb Perspect Biol* **15**, (2023).

77. A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp, G. J. Barton, Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191 (2009).

78. S. R. Eddy, Accelerated Profile HMM Searches. *PLoS Comput Biol* **7**, e1002195 (2011).

79. S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

80. K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066 (2002).

81. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013).

82. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164-1165 (2011).

83. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).

84. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).

85. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528 (2019).

86. M. Ohi, Y. Li, Y. Cheng, T. Walz, Negative staining and image classification - powerful tools in modern electron microscopy. *Biol Proced Online* **6**, 23-34 (2004).

87. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51 (2005).

88. D. Kimanius, L. Dong, G. Sharov, T. Nakane, S. H. W. Scheres, New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem J* **478**, 4169-4185 (2021).

89. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

90. E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023).

91. A. Punjani, D. J. Fleet, 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J Struct Biol* **213**, 107702 (2021).

92. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146-1152 (2019).

93. K. Jamali, L. Kall, R. Zhang, A. Brown, D. Kimanius, S. H. W. Scheres, Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450-457 (2024).

94. T. I. Croll, ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* **74**, 519-530 (2018).

95. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861-877 (2019).

Acknowledgments

We thank Mr. Fuchang Han from Kaifeng City for the schematic diagrams of different species for Figure 1, Yelena Peskova for expert technical assistance, and Drs. Hassane Mchaourab, Itay Budin, and Steven Haddock for helpful discussions. The U-M Cryo-EM Facility has received generous support from the U-M Life Sciences Institute, the U-M Biosciences Initiative, and Beckman Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding:
National Institutes of Health grant R01 GM151635 (AKK)
National Institutes of Health grant R01 HL144131 (AKK, MDO)
National Institutes of Health grant R01GM080403 (JM)
National Institutes of Health grant R01HL122010 (JM)
National Institutes of Health grant R01GM129261 (JM)
National institute of Health grant S10OD030275 (MDO)
National Institute of Health T32GM007315 (SC)
European Research Council's Horizon 2020 grant No. 945026 (DTS)
American Heart Association grant 905705 (SC)
University of Michigan Rackham Predoctoral Fellowship (SC)
Humboldt Professorship of the Alexander von Humboldt Foundation (JM)
HHMI (to Jochen Zimmer)

Author contributions:
Conceptualization: BH, AKK
Methodology: BH, AG, LFLW, SC, DTS, MDO
Investigation: BH, LFLW, SC, DTS
Data curation: BH, SC, LFLW
Formal analysis: BH, LFLW, SC, DTS, MDO
Writing—Original draft: BH, AKK, LFLW, SC, MDO
Writing—Review and editing: BH, LFLW, SC, DTS, AG, JM, EK, AKK, MDO
Visualization: BH, LFLW, SC, DTS
Supervision: AKK, JM, EK, MDO
Project administration: AKK
Funding acquisition: AKK, JM, MDO

Declaration of interests: The authors declare they have no competing interests.

Data availability statement: Maps and models are deposited in the EMDB and PDB (*S. purpuratus* caveolin complex*: EMD-XXXXX, PDB-XXXXX; S. rosetta* caveolin complex*: EMD-XXXXX, PDB-XXXXX)* and will be made available upon publications. All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

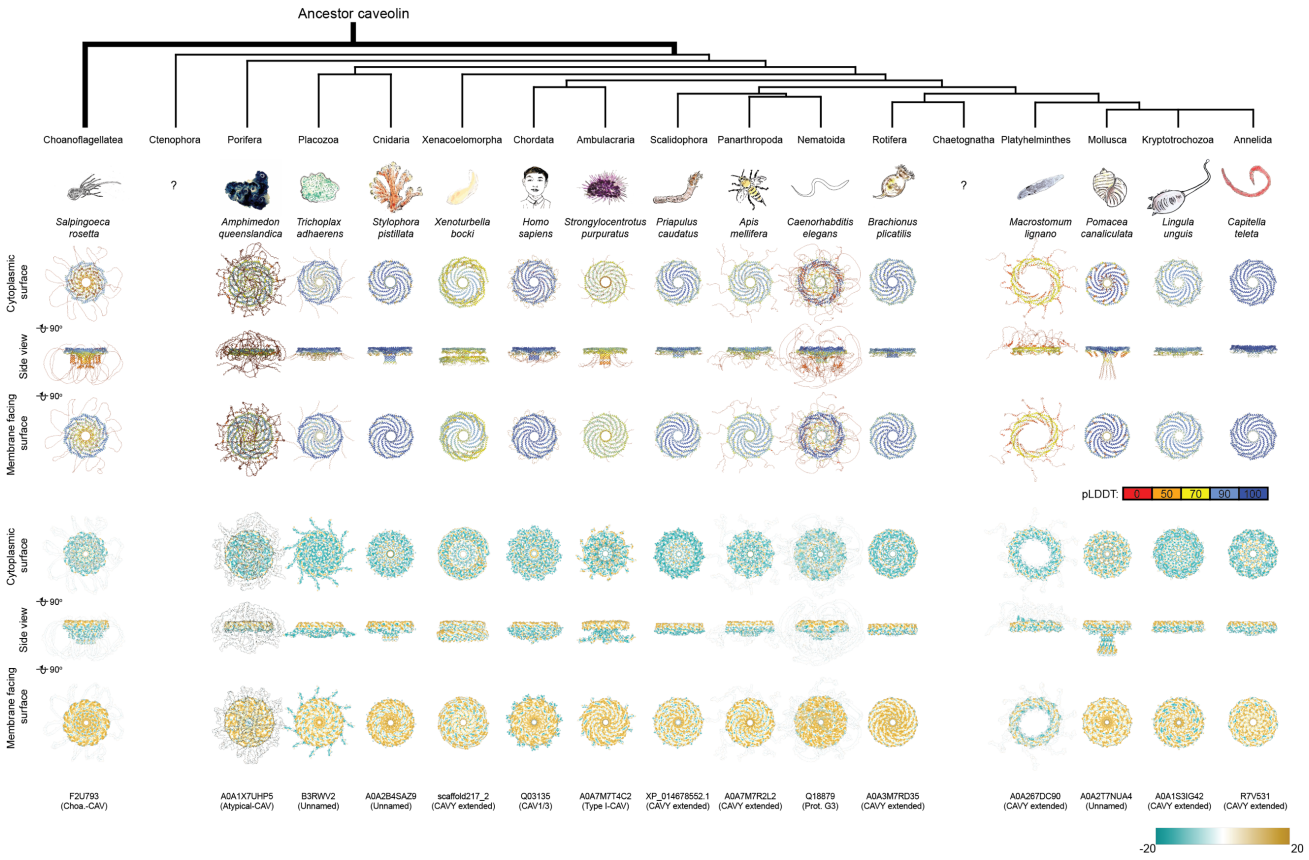

Fig. 1. Conserved structural features of metazoan and choanoflagellate caveolins highlighted by computational modeling predictions. AF2.2 models for 11-mers caveolins from representative species of Choanoflagellatea and fourteen different metazoan phyla/superphyla. Models in top three rows are in ribbon mode and colored by pLDDT confidence values. Models in bottom three rows are in surface mode and colored by lipophilicity values. To better demonstrate the distribution pattern of hydrophobicity in surface mode, we applied transparency to the N-terminus of F2U793, A0A1X7UHP5, A0A7M7R2L2, Q18879, and the residues near the C-terminus of A0A267DC90.

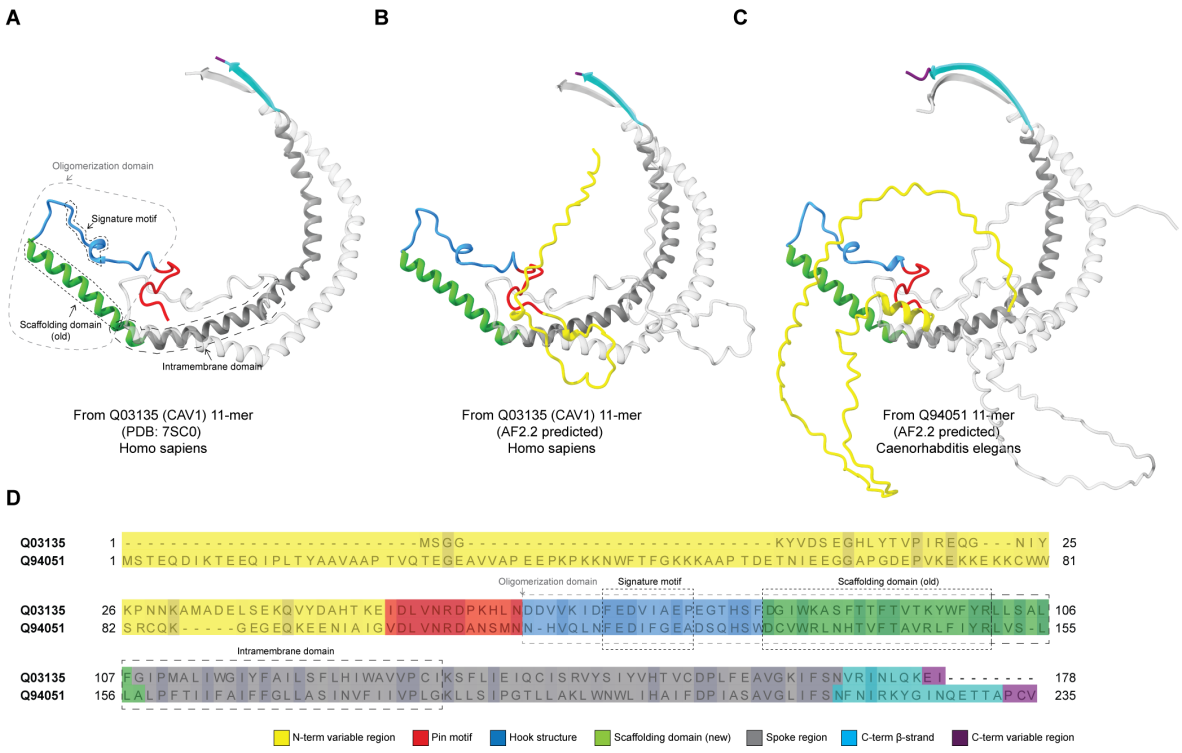

Fig. 2. Proposed structural elements of caveolins. (A-D) Structural elements include a variable N-terminal variable region (yellow), pin motif (red), hook region (blue), scaffolding domain (green), spoke region (blue), β-strand (cyan), and C-terminal variable region (purple). For illustration purposes, elements are mapped into the structures of **(A)** two neighboring protomers from the cryo-EM based secondary structure model of human Cav1 8S complex (PDB: 7SC0), **(B)** two neighboring protomers from the 11-mer of human Cav1 (Q03135) predicted by AF2.2, and **(C)** two neighboring protomers from the 11-mer of *Caenorhabditis elegans* caveolin (Q94051) predicted by AF2.2. The comparison between the old domain segregation and our proposed segmentation using structural elements can be observed in both the model panel (A) and sequence alignments of human Cav1 and *C. elegans* caveolin Q94051 **(D)**.

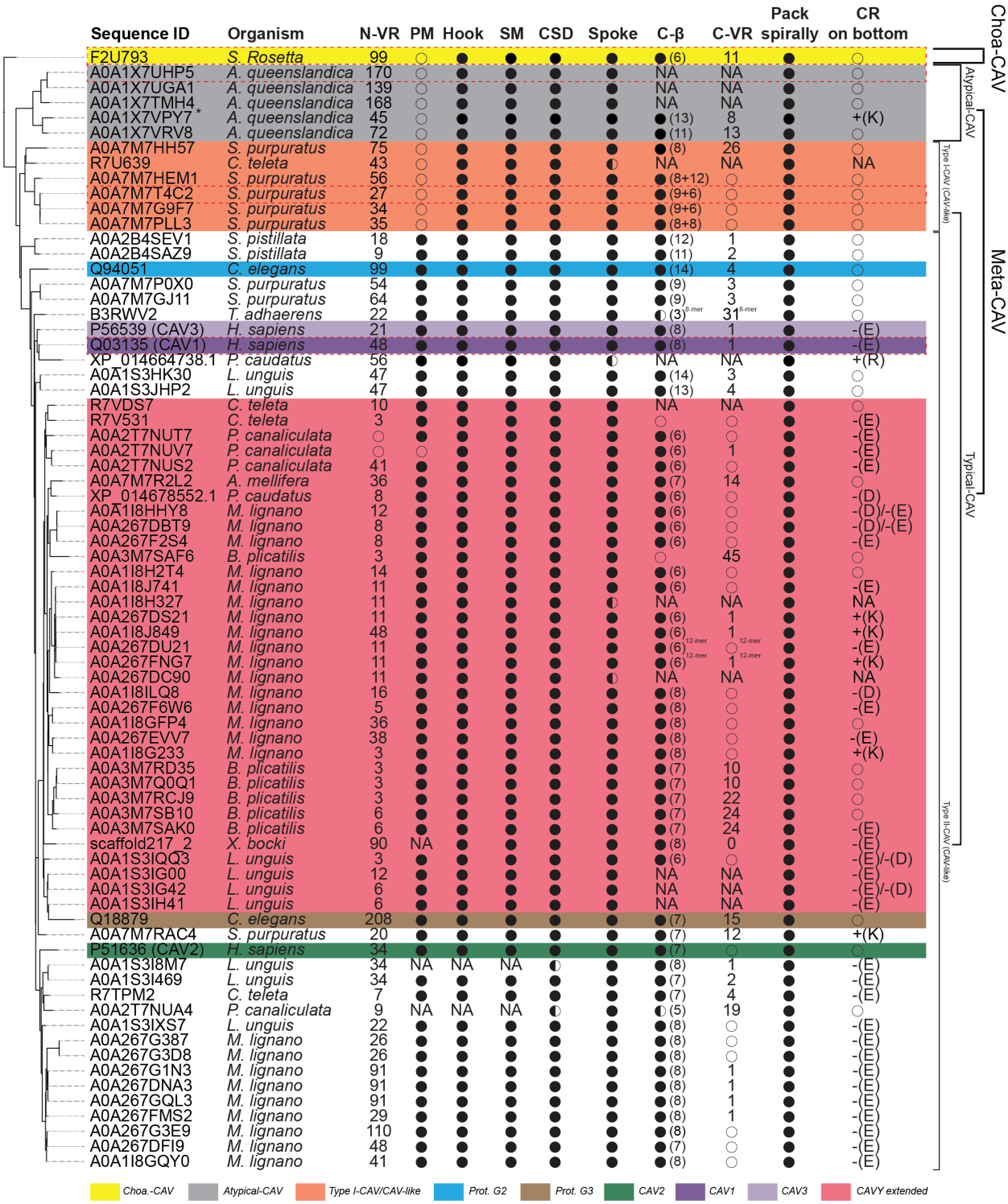

Fig. 3. Summary of the structural features of caveolins suggested by computational modeling predictions. Phylogenetic tree shown on the left-hand side of the table is based on Figure S1. N-VR; N-terminus variable region; PM, pin motif; Hook, hook structure; SM, signature motif; CSD, scaffolding domain; Spoke, spoke region; C-β, C-terminus β-strand; C-VR, C-

terminus variable region; Pack spirally, whether the protein sequence is predicted to be assembled into a disc-shaped oligomer; CR on bottom, charged residues on hydrophobic side of complexes ("+" represents a positive charge, "-" represents a negative charge, and uppercase English letters represent the corresponding amino acid residue abbreviations). Structural features were summarized mainly based on AF2.2 predicted 11-mers unless otherwise noted in the upper right corner of particular features. ●, structural feature was predicted to be present; ◑, structural feature was predicted to be partially present; ○, structural feature was not predicted to be present. NA, not applicable (corresponding sequences were either missing or shifted due to sequence missing). *, A0A1X7VPY7 was not predicted by AF2.2 to form any homologous disc-like or ring-like complexes. However, it was predicted to form a hybrid complex structure with A0A1X7VRV8. The structural features listed for A0A1X7VPY7 are summarized from the model of the hybrid complex consisting of A0A1X7VPY7 (1 copy) and A0A1X7VRV8 (10 copies).

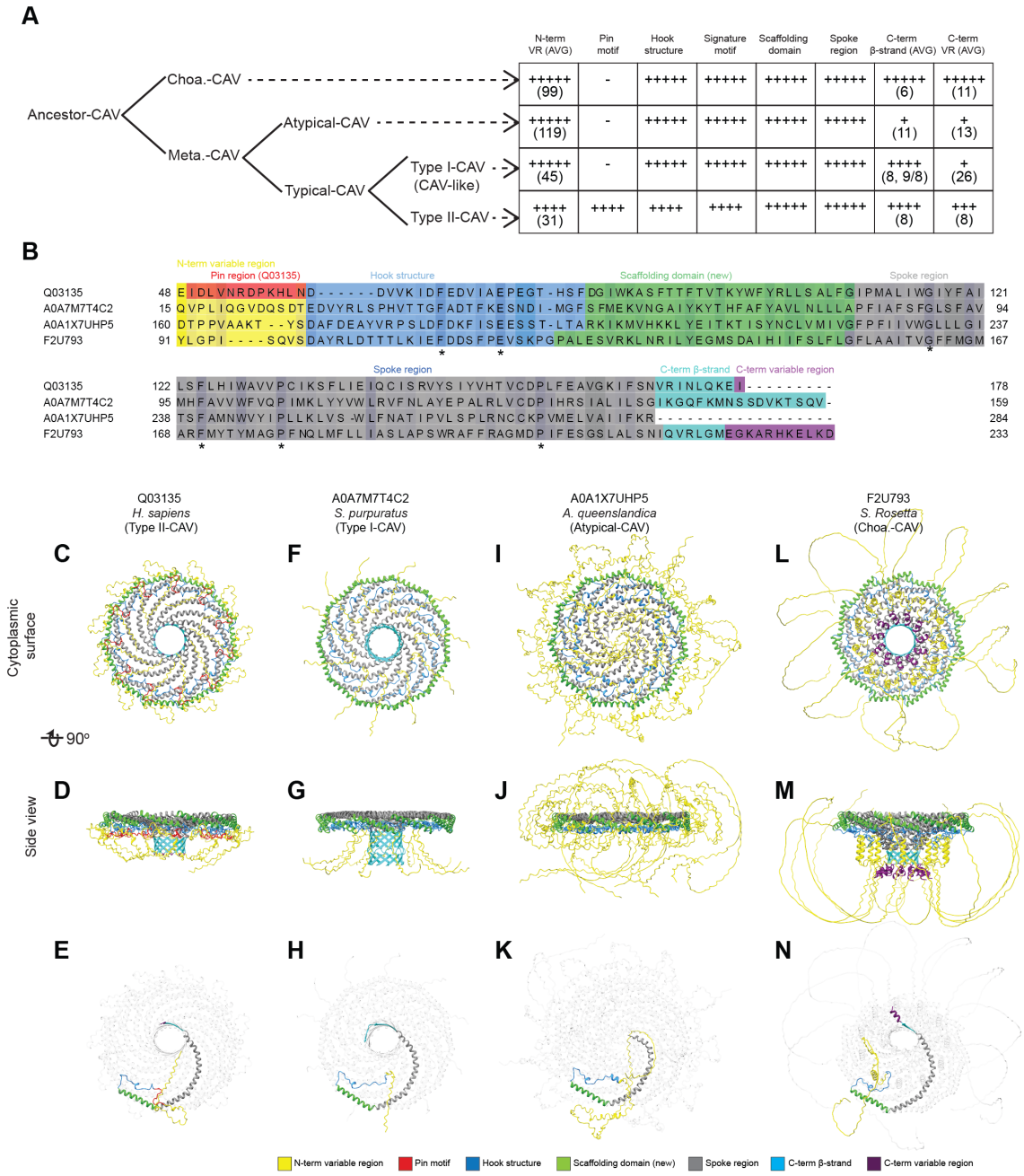

Fig. 4. Predicted structural features of representative metazoan and non-metazoan caveolins. (A) Model summarizing the key structural similarities and difference in different groups of caveolins based on the phylogenetic analysis and structural comparisons presented in Table 1 and Supplemental File 2. **(B)** Sequence alignment of representative caveolins. Conserved residues are colored with different intensity according to the percent identity. The range of different structure regions were also highlighted with different colors. **(C-N)** Different views of AF2.2 models of representative caveolins**. (C-E)** Human Cav1 (Type II-CAV, Q03135). **(F-H)** *S. purpuratus* (Type I-CAV, A0A7M7T4C2). **(I-K)** *A. queenslandica* (Atypical-CAV, A0A1X7UHP5). **(L-N)** *S. rosetta* (Choa-CAV, F2U793). The range of different structure regions were also highlighted with different colors. To better display the structure of a single protomer within the complex, the ten protomers in the models in panels D, G, J, and M have been made transparent.

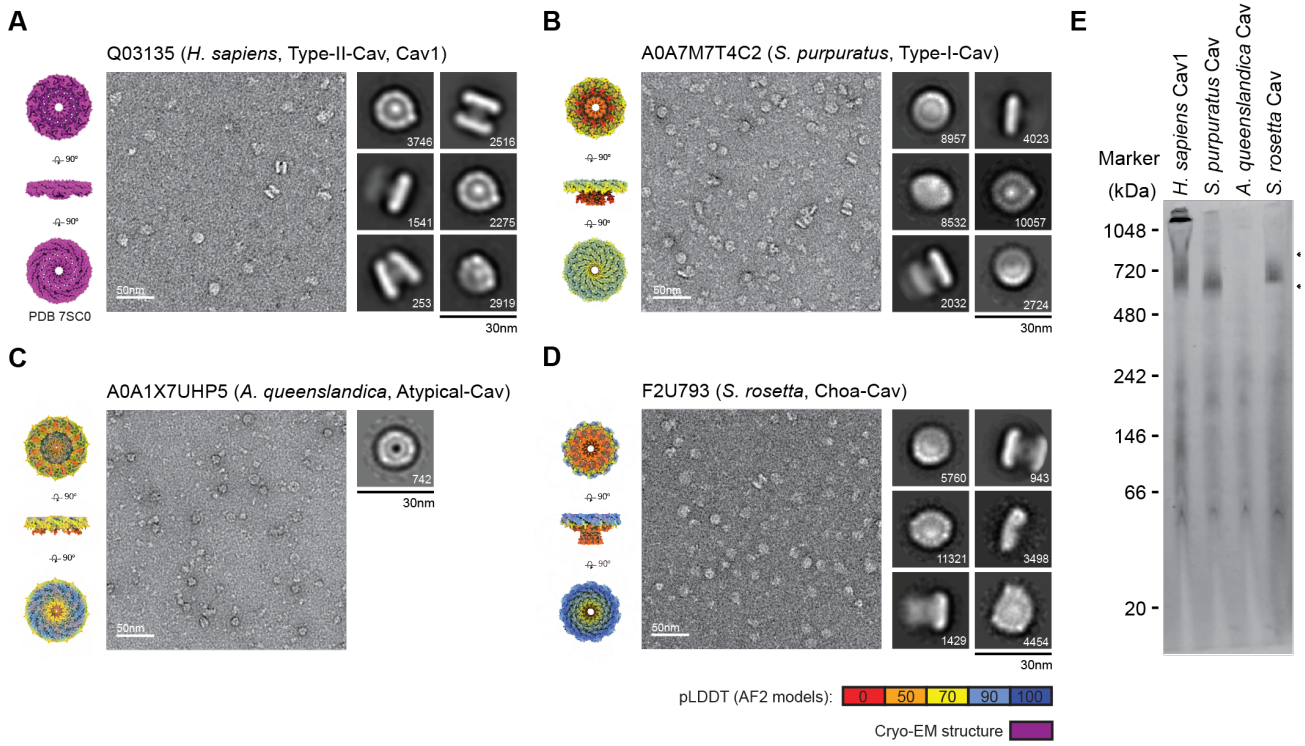

Fig. 5. Negative stain EM shows diverse caveolins assemble into disc-shaped complexes. **(A)** *H. sapiens* Cav1, (**B**) *S. purpuratus* caveolin, (**C**) *A. queenslandica* caveolin, and (**D**) *S. rosetta* caveolin. (**A-D**) In each panel, surface-filling models for the cryo-EM structure or AF2.2 11-mer structures are shown on the left, representative images of negatively stained caveolin complexes are shown in the middle, and representative 2D averages of caveolins are shown on the right. The number of particles found in each class average is shown in the bottom right. Scale bar, 30 nm. (**E**) Blue Native gel with arrows highlighting the position of 8S-like complexes. MW markers are indicated.

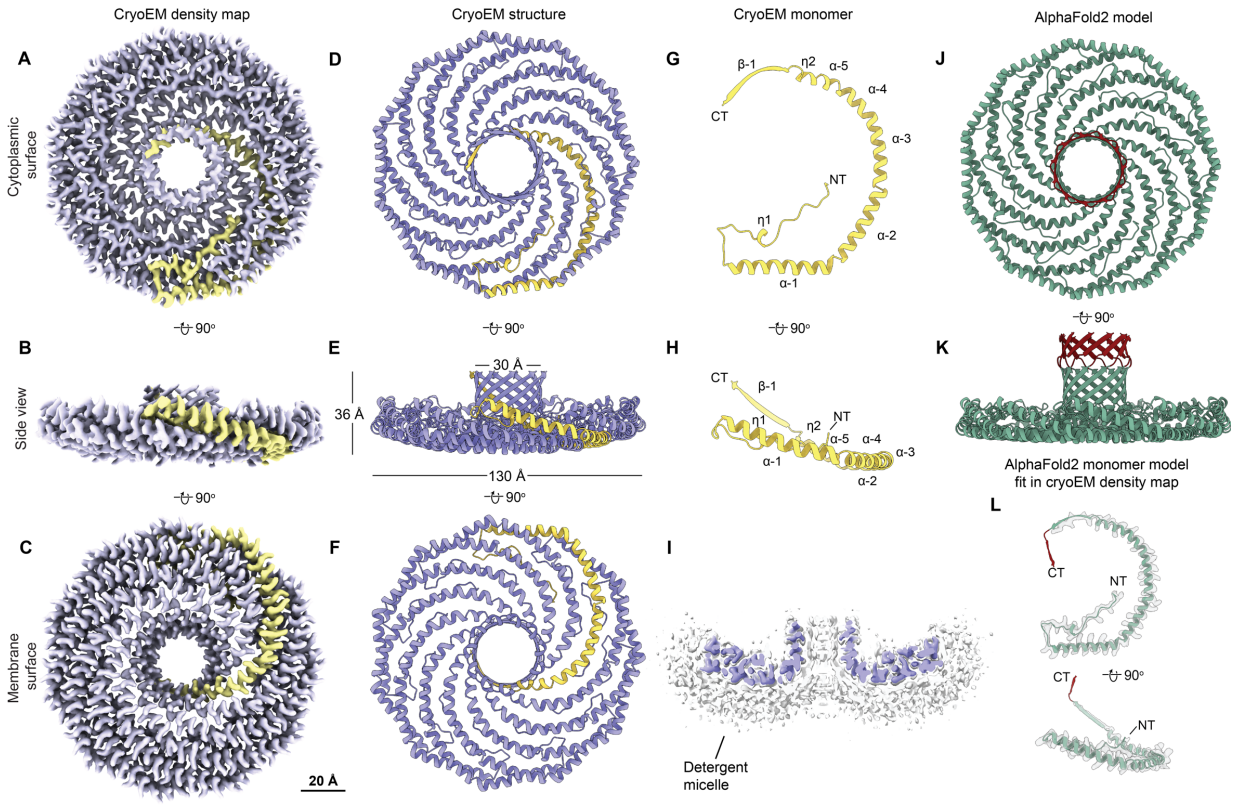

Fig. 6. *S. purpuratus* caveolin forms an 11-mer complex. (A-C) A 3.1 Å resolution cryo-EM density map of the *S. purpuratus* Caveolincomplex with 11-fold symmetry. Complex is shown with ninety-degree rotated views displaying the cytoplasmic-facing surface (**A**), side (**B**), and membrane-facing surface (**C**). The complex has an overall disc-like structure with 11 spiraling α-helices and a central β-barrel. A single protomer is highlighted in yellow. Scale bar, 20 Å. (**D to F**) Secondary structure model of the *S. purpuratus* caveolin in the same views as shown in panels A-C. (**G and H**) Secondary structure of *S. purpuratus* caveolin protomer with secondary features and N- and C-termini noted. (**I**) Central slice of the density map (purple) with the detergent micelle (grey). (**J, K**) AF2.2-predicted structure of the *S. purpuratus* caveolin 11-mer showing views of the cytoplasmic-facing surface (J) and side view (K). Structured regions predicted by AF2.2 that are not found in the cryo-EM structure are highlighted in burgundy. (**L**) Protomer from the AF2.2 11-mer model fit into the cryo-EM density map (gray outline) of a protomer.

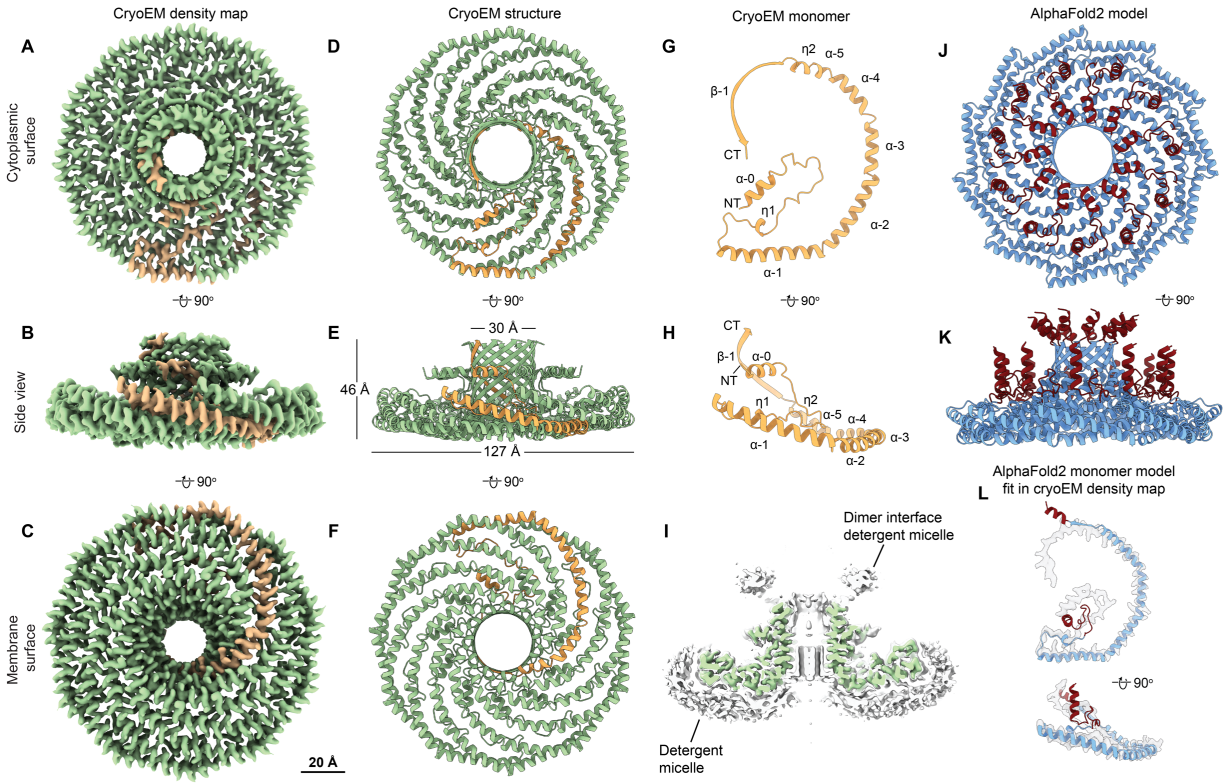

Fig. 7. *S. rosetta* caveolin forms an 11-mer complex with elongated β-barrel and extended N-terminal region. (**A-C**) A 2.9 Å resolution cryo-EM density of *S. rosetta* caveolin complex with 11-fold symmetry. The complex is shown with ninety-degree rotated views displaying the cytoplasmic-facing surface (**A**), side (**B**), and membrane-facing surface (**C**). The complex has an overall disc-like structure with 11 spiraling α-helices, a central elongated β-barrel, and extended N-terminal region. A single protomer is highlighted in orange. Scale bar, 20 Å. (**D to F**) Secondary structure model of the *S. rosetta* caveolin complex in the same views as shown in panels A-C. (**G and H**) Secondary structure of *S. rosetta* caveolin protomer with secondary features and N- and C-termini noted. (**I**) Central slice of the density map (green) with the detergent micelle (grey). (**J, K**) AF2.2-predicted structure of the *S. rosetta* caveolin 11-mer showing views of the cytoplasmic-facing surface (J) and side view (K). Structured regions predicted by AF2.2 that are not found in the cryo-EM structure are highlighted in burgundy. (**L**) Protomer from the AF2.2 11-mer model fit into the cryo-EM density map (gray outline) of a protomer.

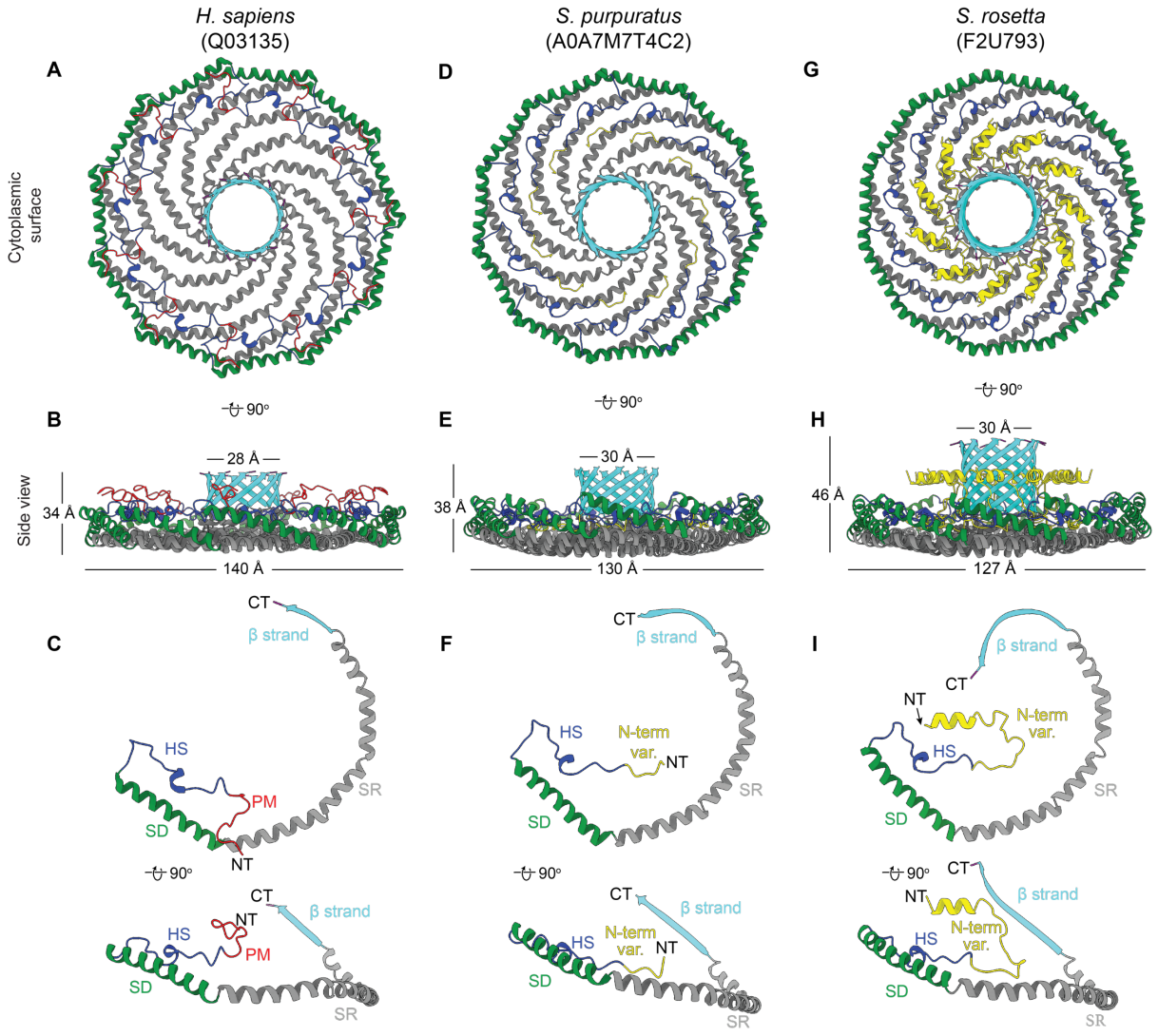

Fig. 8. Comparison of cryo-EM structures for *H. sapiens*, *S. rosetta,* and *S. purpuratus* caveolin complexes. (A, B, D, E) Cytoplasmic-facing surface and side views of cryo-EM structures are shown for human Cav1 (**A, B**), *S. purpuratus* caveolin (**D, E**), and *S. rosetta* caveolin (**G, H**). (**C, F, I**) Protomer structures extracted from the 11-mers are shown in same views as panel above. Domains are colored as follows: N-terminal variable region (yellow), pin motif (PM, red), hook structure (HS, blue), scaffolding domain (SD, green), spoke region (SR, gray), β-strand (cyan). NT, N-terminus; C-term, C-terminus.

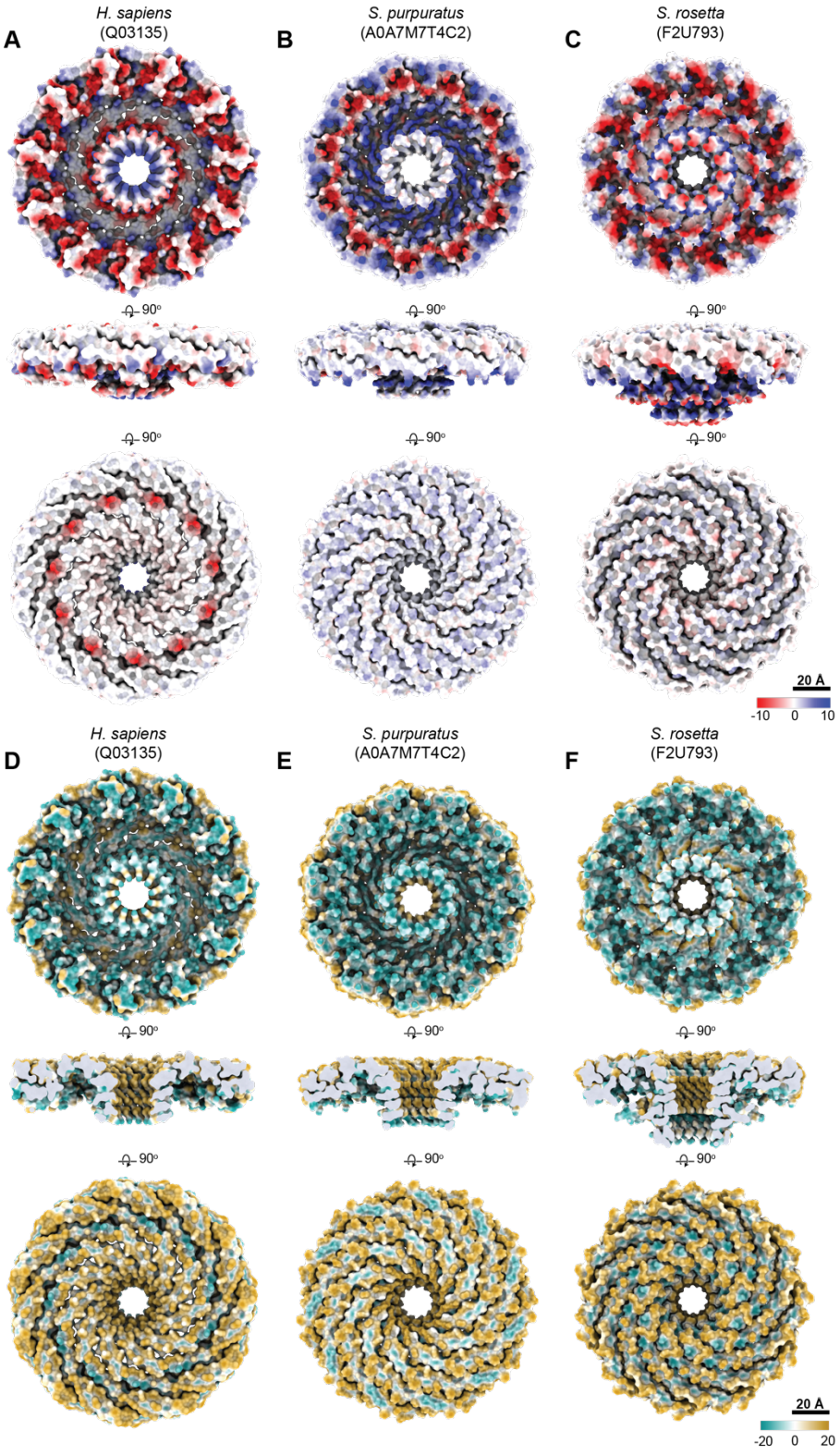

Fig. 9. Distribution of charged residues and hydrophobicity of the predicted membrane and cytoplasmic facing surfaces of *H. sapiens*, *S. rosetta,* and *S. purpuratus* caveolin complexes. (A-C) Space filling models of the caveolin complexes rotated 90˚, showing (**A-C**)

the charge of the amino acids or (**D-F**) hydrophobicity values. Note that side views in A-C are shown with the surface of the complex whereas a cut through the center of the complex is shown in D-F**.**

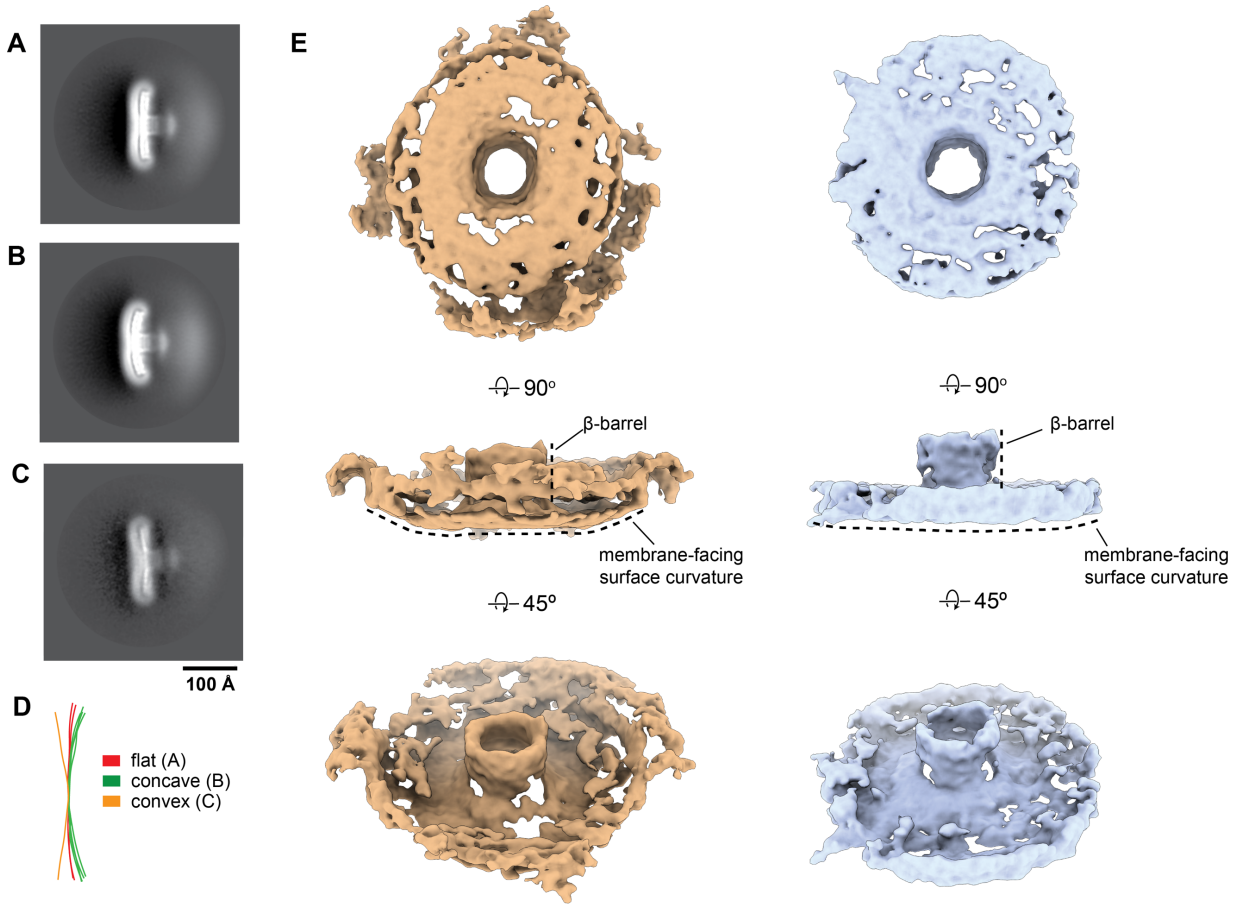

Fig. 10. *S. purpuratus* caveolin displays various curvatures in 2D classes and 3D variability analysis. (**A-C**) *S. purpuratus* caveolin 2D class averages that show a flat (**A**), concave (**B**), and convex (**C**) curvature of the complex. Scale bar, 100 Å. (**D**) The difference in curvature is highlighted with overlaid traces of the membrane-facing surface from 2D class averages. (**E**) Structures representing the negative and positive values along the reaction coordinate of a 3D variability analysis (3DVA) component calculated with a resolution limit of 8 Å. The complex on the left (orange) shows a concave membrane surface curvature while the complex on the right (light blue) shows a flat membrane surface curvature and "lifting" of the β-barrel above the rim of the complex. The proposed membrane facing surface is shown (top), rotated 90˚ around the x-axis to show a side view (middle), and rotated an additional 45˚ to show a view of the predicted cytoplasmic facing surface (bottom).