Design principles of caveolins across metazoa and beyond

Bing Han [1,2], Louis F. L. Wilson [2,3], Alican Gulsevin [4, 5], Jens Meiler [4,6], Erkan Karakas [7], 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] Howard Hughes Medical Institute, University of Virginia School of Medicine, Charlottesville, VA, USA
[4] Department of Chemistry, Vanderbilt University Nashville, TN, USA
[5] Center for Structural Biology, Vanderbilt University, Nashville, TN, USA
[6] Institute for Drug Discovery, Leipzig University, Germany
[7] Department of Molecular Physiology and Biophysics, Vanderbilt University Nashville, TN, USA

*Correspondence: akk7hp@virginia.edu

Short title: Caveolins across evolution

Abstract

Caveolins are a unique family of membrane remodeling proteins essential for building flask-shaped nanoscale plasma membrane invaginations that regulate signaling and stress responses in vertebrates known as caveolae. Recent evidence suggests that to generate caveolae, human caveolin-1 assembles into amphipathic disc-shaped complexes consisting of 11 tightly packed, spirally organized protomers. Interestingly, caveolins are expressed broadly throughout Metazoa, including in organisms that lack conventional caveolae. The structural features of caveolins that control their functionality in diverse organisms have yet to be established. To address this fundamental question, we use AlphaFold2 as a predictive tool to investigate the structural properties of 72 representative caveolins across evolution. The results of this analysis reveal caveolins consist of six common structural elements that can be mixed and matched, uncover principles that govern their assembly into oligomeric discs, and provide evidence that the association of caveolins with membranes is one of the most ancient functions of the protein. We also identify homologs of caveolins in choanoflagellates, pointing to a pre-metazoan origin of the protein family. Together, these findings define a new structure-based framework for probing the functional roles of caveolins across evolutionary space.

Main text

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 multicellular animals, a dramatic expansion in membrane-associated proteins occurred [3]. Among the membrane proteins thought to have first emerged in Metazoa is the caveolin family of membrane remodeling proteins [3-5]. Best recognized for their role in vertebrates as structural components of flask-shaped invaginations of the plasma membrane known as caveolae, caveolins have been identified in most metazoan clades, suggesting they fulfill essential roles in multiple organisms [4-6]. Caveolins and caveolae are distributed throughout the body and serve as important regulators of multiple organ systems [7-10]. Furthermore, caveolins and caveolae have been more broadly implicated in regulation of cell signaling, lipid metabolism, and sensing and responding to stress[9,11-14].

Unlike traditional vesicle coat proteins such as clathrin, COPI, and COPII that cycle on and off membranes and share evolutionary origins and structural features [2], caveolins are unrelated in sequence to other proteins and remain integrated in membranes throughout their life cycle. Caveolins lack a traditional transmembrane domain and are instead embedded within a single leaflet of cell membranes, with both their N- and C-termini facing the cytoplasm [15-17]. In humans, the caveolin gene family includes three family members: CAV1, CAV2, and CAV3 [9,18]. Expression of CAV1 and CAV3 is essential for caveolae biogenesis in non-muscle and muscle cells, respectively [19-27]. Both CAV1 and CAV3 function as highly stable oligomers containing between 7 and 14 protomers [28-33]. Oligomerization occurs early in the secretory pathway and is required for their ultimate delivery to the plasma membrane [34,35]. Although itself incapable of supporting caveolae assembly in the absence of CAV1, CAV2 has a number of important physiological roles [36-41].

Most studies of caveolins have focused on their roles in mammalian cells where caveolae are often abundant [16,42,43]. However, caveolins can function independently of caveolae in some cell types lacking a second class of proteins required for caveolae biogenesis known as cavins [44-47]. Furthermore, the expression of members of the cavin family is thought to be limited to vertebrates [48]. 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 [5,49-52]. 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 human caveolin-1 (CAV1) homo-oligomeric 8S complex using cryo-electron microscopy (cryo-EM) [53,54]. The complex is composed of 11 copies of CAV1 that are symmetrically arranged to form a disc-like structure. N-terminal sequences located on the outer rim of the disc generate a “key-lock” structure consisting of a pin motif and oligomerization domain that play a critical role in interlocking adjacent protomers together to maintain the complexes disc or ring-like shape. From the N-terminal domains, a long amphipathic α-helix from each protomer extends toward the center of the disc. A parallel β-barrel formed from single β-strands contributed by the C-terminal-most residues of each monomer further stabilizes the complex. Yet even in light of this major advance, caveolins remain among the most enigmatic and elusive membrane proteins in biology, and the relationship between the structure and function of caveolins, and how this differs across family members and species, is still largely unknown.

In the past year, tremendous advances have been made in structural elucidation through the development of deep learning-based methods such as AlphaFold2 (AF2), capable of predicting protein structures based on their amino acid sequence [55]. We recently showed that despite the unusual feature of human CAV1, AF2 is able to predict its overall fold and reproduce many key structural features of the CAV1 8S complex [56]. In the current work, we now use AF2 to investigate the structural features of caveolins across evolution.

Results

AF2 can explain several experimental features of human CAV1, CAV2, and CAV3, making it a useful tool for studying the structural properties of caveolins more broadly

As a starting point for our analysis, we compared the structures of human CAV1, CAV2, and CAV3. Human CAV2 is 38% identical and 58% similar to human CAV1 [37]. Human CAV3 lacks the first 27 amino acids found in CAV1 but is otherwise 65% identical and 85% similar to Cav1 [25,26] (**Figure 1A**). To understand how these differences in sequence translate to differences in structure, we first obtained the predicted structures of human CAV1, CAV2, and CAV3 monomers from the AF2 database [57] and compared them to the Cryo-EM structure of the human CAV1 complex (**Figure 1)**. As for the case of human CAV1, CAV2 and CAV3 are predicted to contain extended helical regions with a high confidence level, whereas the N- and C-terminal regions consist primarily of coils predicted with medium to low confidence (**Figure 1**).

To better compare the structural features of the protomers, we mapped the position of regions of CAV1 identified as structurally or functionally important in the literature or in light of our analyses based on the recent cryoEM structure. These features include the pin motif, a series of residues in the outer rim of the complex that locks adjacent protomers into place [53]; the

signature motif, the most highly conserved region of the protein [26,37]; the scaffolding domain, a region important for both oligomerization as well as serving as a protein-protein interaction interface [58,59]; the intramembrane domain, which the structure reveals tightly packs with neighboring protomers to form a flat membrane-binding surface [53]; the spoke region, consisting of amphipathic α-helices immediately following the intramembrane domain [53]; and a C-terminal β-strand that interacts with other protomers to form a parallel β-barrel with a hydrophobic interior [53]. This analysis reveals that the predicted structure of both the CAV2 and CAV3 monomers share many features with that of CAV1, including the presence of a loop at the N-terminal region that makes a 180° turn and the helical nature of the scaffolding domain, intramembrane domain, and spoke region (**Figure 1D- I**). One major difference is that, unlike CAV1 and CAV3, the CAV2 monomer does not bend between helix 1 (green) and helix 2 (purple) (**Figure 1F and G**).

Next, we compared the predicted structures of homo-oligomers of human CAV1, CAV2, and CAV3 using AF2.1 (**Figure S1**) and AF2.2 (**Figure 2**). Both AF2.1 and AF2.2 are capable of assembling homo-oligomeric complexes. AF2.1.0 includes the original AlphaFold-Multimer data pipeline, model, and metrics. AF2.2 incorporates updated AlphaFold-Multimer model parameters [60] and is thought to reduce the number of clashes on average and be slightly more accurate than AF2.1. It also allows for the evaluation of larger complexes due to the higher upper limit of residues that can be analyzed.

As recently reported [56], both AF2.1 and AF2.2 predict that CAV1 forms disc-shaped complexes assembled from spirally packed long helices. The complexes also contain a characteristic pin motif and a central β-barrel formed by strands contributed by the C-termini of adjacent protomers (**Figure 2A-C, Figure S1A-C**). Multiple copies of CAV3 are also predicted to assemble into a disc-shaped complex similar in overall structure to that observed experimentally for CAV1 (**Figure 2G-I, Figure S1G-I**). This agrees with biochemical and structural evidence that both CAV1 and CAV3 form 8S complexes similar in overall shape and size [33,61]. One notable distinction is that the N-terminal most residues of CAV3 are predicted to form a helical structure, unlike that of CAV1, which are primarily disordered. The presence of this predicted α-helix in CAV3 observation agrees with a preliminary report from another group [62].

Interestingly, CAV2 was also predicted to form multimers containing between 2 and 11 protomers (**Figure 2D-F, Figure S1D-F**). However, the CAV2 oligomers either fail to generate a closed disc-like structure or form discs that are highly asymmetric and twisted (**Figure 2D-F, Figure S1D-F**). These findings suggest that differences in angles between helix 1 and helix 2 of CAV1/CAV3 and CAV2 protomers may play a critical role in determining whether the protomers pack together into tightly apposed spirals that ultimately form a closed disc. It is also consistent with observations that unless CAV1 is present, CAV2 fails to assemble into high molecular weight oligomers or support caveolae assembly [34,63-65]. We did however note that when we repeated these predictions at later dates, AF2.2 showed an increased tendency to assemble CAV2 into more regular, symmetric complexes (**Figure S2**).

In summary, both available versions of AF2 predict the overall secondary, tertiary, and quaternary structure of CAV1. They also assemble CAV1 protomers into a disc-shaped complex that shares many features with the experimentally determined 8S complex structure and predict that CAV3 can also form closed, symmetric disc-shaped complexes but that CAV2 cannot. These findings suggest AF2 can be a useful tool to propose testable hypotheses on both structure and function of caveolins. Given that the only high-resolution experimental structure of any caveolin family member currently available is for human CAV1, the relative

accuracy of AF2.2 compared to AF2.1 in predicting other structural forms of caveolins remains unknown. In the following analysis, we, therefore, employed AF2.1 to predict the structures of representative caveolins across Metazoa and then used the prediction results of AF2.2 to cross-validate and supplement these findings.

The formation of helically-based spiral multimers with a hydrophobic membrane facing surface is a conserved feature of metazoan caveolins

The above analysis suggests caveolins capable of supporting caveolae biogenesis form closed, regular, disc-shaped oligomeric complexes. This raises the fundamental questions of when mammalian caveolins gained these behaviors and which if any of its current structural features were directly inherited from the ancestral caveolin. The degree of structural and functional diversity of caveolins across species is also currently unknown. To address these questions, we performed large-scale structural predictions on 72 examples of PFAM-annotated caveolin family members from representative species of 13 different Metazoan phyla/superphyla (**Supplemental File 1**). For our initial analysis, we generated *n*-mers of increasing size with AF2.1 (**Supplementary File 1**). Structures with interesting properties were subsequently analyzed using AF2.2.

The results of this analysis reveal that packing of caveolins in a spiral pattern is a common feature for most of the caveolins examined. Indeed, caveolins found in almost all species we examined are predicted to form closed spiral-packed discs (**Figure 3, Figure S3, and Supplemental File 1**). Several of these structures also were predicted to contain parallel β-barrels in the center, similar to the structure of human CAV1. This suggests that even in the most distantly related metazoans, at least a portion of caveolins oligomerize like their counterparts in other phyla/superphyla. A notable exception is the sequences from *Amphimedon queenslandica,* a sponge native to the Great Barrier Reef. As one of the most distantly related metazoans to humans, this species is a useful model to examine the evolution of metazoans [66]. None of its caveolins were predicted to form closed discs by AF2.1 (**Figure S3**). However, most of them still tend to organize into spiral-like assemblies (**Figure S3 and Supplemental File 1**). Interestingly, AF2.2 predicted most *Amphimedon queenslandica* caveolins to form closed disc or ring-like structures with median to low confidences (**Figure 3**).

Mammalian caveolins are monotopic membrane proteins [15-17]. Consistent with this, the disc-shaped human CAV1 complex is primarily hydrophobic both on its membrane facing surface and around the outer rim. To test if other caveolins share this feature, we compared the distribution of hydrophobic and hydrophilic residues on the predicted structures. In all models examined, the plane formed by the long helices is predicted to be amphipathic, consisting of a primarily hydrophobic membrane-facing surface and a hydrophilic surface likely facing the cytoplasm (**Figure 3, Figure S3**). This amphipathic feature extends to include the outer rims of the complexes, as well as the interior of the C-terminal β-barrels for caveolins that contain these structural elements. The association of caveolins with membranes may thus be an ancient and highly conserved feature of the protein. Around 45% of the 72 Cav1 complexes we examined have a completely hydrophobic membrane facing surface, whereas others, including human CAV1, contained one or more charged residues positioned in different patterns (**Figure S4**). The presence of these charged residues in the membrane facing surface suggests they may be functionally important, but their biological roles remain to be determined.

Caveolins consist of six basic structural units

We next sought to identify specific structural motifs found in multiple caveolins across evolution. Previous studies have identified the oligomerization, scaffolding, and intramembrane domains as functionally important regions of mammalian caveolins [15,17,30,58,59] . However, 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 [53] **(Figure 1**). We thus instead used the structure itself as a template to deconstruct the protein into six structurally defined domains (**Figure 4**). To illustrate their position in the structure and relationships to one another, we mapped them onto the predicted structures of two different caveolins: human CAV1 and one of the two *C. elegans* caveolins (Q94051) (**Figure 4**). From the N-terminus to the C-terminus, these six basic structural units are as follows:

1) **N-terminal variable region**. In human CAV1, residues 1-48 are predicted to be disordered and were not resolved in the cryoEM structure [53]. Similarly, other caveolins are predicted to contain N-terminal disordered regions. Both the length and sequence of the most N-terminal region are highly variable across caveolins. For example, for *C. elegans* caveolin Q94051, this region is predicted to be 99 residues in length (**Figure 4C, in red**). The structure of this region is usually predicted by AF2 with low confidence, further illustrating its flexible nature (**Figure 1, Supplemental File 1**).

2) **Pin motif.** First identified in the experimental structure of the human CAV1 8S complex [53], the pin motif (residues 49-60 in human CAV1) is critical for neighboring protomers to interact. It stacks on top of the neighboring protomer in the complex, stabilizing the protomer-protomer interaction at the rim portion of the disc-shaped 8S complex [53]. A similar motif is predicted to exist in *C. elegans* caveolin Q94051 (**Figure 4C, in yellow**).

3) **Hook structure.** Residues 61-81 of human CAV1 consist of a loop that undergoes a 180° turn (**Figure 4A-B, in green**). It corresponds to the first half (residues 61-81) of the oligomerization domain (residues 61-101) of human CAV1. Embedded within this same region is the highly conserved signature motif (residues 68-75) (**Figure S5**). In the cryo-EM structure, the signature motif consists of a 3/10 helix followed by a short coiled structure (**Figure S5**). The positioning of the signature motif enables it to form hydrogen bonds with nearby residues that help to sustain the hook structure as well as interactions with neighboring protomers (**Figure S5**).

4) **Bow-like long helical region.** Residues 82-169 of human CAV1 consist of a series of α-helices connected in tandem by short loops, generating a long helical region in the middle of the protein. All turns and kinks occur along the same orientation and plane, bending the helices into an arc of about 180°. This region of the protein encompasses the scaffolding domain (residues 82-101), the intramembrane domain (residues 102-134), and the spoke region (residues 135-169) (**Figure 4A-B, in cyan**). The presence of these bow-like helices is a characteristic feature of most caveolins examined (**Supplemental File 1**).

5) **C-terminal β-sheet.** Following the helical region, the C-terminal domain of each human CAV1 protomer folds into a β-strand (**Figure 4A-B, in orange**). Together with neighboring protomers, the strands assemble into a parallel β-barrel. In human CAV1, the strand is 7 residues long (residues 170-176). In other caveolins examined here, the predicted strands range in length from 3 to 18 residues.

6) **C-terminal variable region.** Human CAV1 has only two residues immediately following the β-strand at the C-terminus. A subset of other caveolins also contains a similar region (C-terminal variable region) with differences in length and composition across caveolins (**Figure 4, in purple**). The structure of this region is typically predicted by AF2 with low confidence, suggesting it is disordered (**Supplemental File 1**).

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

Previous studies have suggested that caveolins are only found in metazoans [3-5]. 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 [5]. Whereas CavY appears to have been lost in most vertebrates, CavX seems to have given rise to Cav1 and Cav3, the 'canonical' caveola-forming family members [5]. In a similar fashion, CavZ appears to have given rise to the Cav2 and Cav2R groups, Cav2R having later been lost from placental mammals [5]. Other clades include 'Group 1' protostome sequences that appear to form a monophyletic group with CavY descendents ('CavY extended'), Protostomia Groups 2 and 3, and a more distantly related group coined the 'caveolin-like' group [5]. However, the structure of the caveolin family's common ancestor, as well as the similarities and differences in structure across clades are currently unknown.

To address this question, we first performed an updated phylogenetic analysis. To do so, we inferred a maximum-likelihood phylogeny using a set of previously categorized caveolin and caveolin-like sequences [5] together with 72 protein sequences with caveolin PFAM annotations from 13 distantly related holozoans (**Table S1**). Two of these species, *Amphimedon queenslandica* (Porifera, Metazoa) and *Salpingoeca rosetta* (Choanoflagellatea, the sister group to Metazoa), are of particular evolutionary interest due to their early divergence with respect to vertebrates [66,67]. Caveolins are absent from the genome of the choanoflagellate *Monosiga brevicollis* and had previously been described as exclusive to metazoans [3-5]. However, a caveolin-like protein has been identified in *Monosiga ovata* [68]. To help confirm the homology of the *S. rosetta* sequence (UniProt: F2U793) and the *A. queenslandica* sequences with previously characterized caveolins, we built a caveolin profile from the latter using HMMER. Searches against the *S. rosetta* and *A. queenslandica* proteomes retrieved the relevant sequences with *E*-values of 2.5×10^{-11}, indicating a confident prediction of homology. These findings suggest that caveolins are in fact not limited to Metazoa, but are also found in the closest living relatives of Metazoa [67]. The caveolin sequence from *S. rosetta* was thus used as a tentative outgroup for our analysis.

Though achieving only weak 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 (**Figure 5**). Many of the 72 sequences of interest could be assigned to one of these previously defined groups (**Figure 5**). As expected, the caveolin sequence from *S. rosetta* (F2U793) was separated from all other metazoan caveolins on the phylogenetic tree (**Figure 5**, **Table 1**). We named the clade this sequence belongs to as Choa-CAV. Correspondingly, we named all metazoan caveolins as Meta-CAV (**Table 1**). Within the Meta-CAV group, the selected caveolin homologs from *Amphimedeon 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 of the relatively well studied caveolins belong to Type II-CAV, such as human CAV1, CAV2, and CAV3, as well as *C. elegans* caveolins and *Apis mellifera* (honey bee) caveolin. Type 1-CAV corresponds to the CAV-like clade in a previous study [5].

Comparison of structural features of metazoan and non-metazoan caveolins

Next, we compared the structural features of caveolins across clades. For this analysis, we used AF2 to predict the structures of 72 representative caveolins selected based on their position in the phylogenetic tree (**Table 1**). We compared how many of the six basic structural units each caveolin contained, their propensity to form spiral assemblies, and the distribution of charges on their membrane-facing surface. To better illustrate key similarities and differences across caveolins, caveolin-like proteins (Type 1-CAV), and their ancestors, we selected representative examples from major classes of caveolins, including the choanoflagellate caveolin from *S. rosetta* (F2U793), an example of an *A. queenslandica* caveolin (A0A1X7UHP5) from the newly identified clade, and a "CAV-like (Type I-CAV)" protein from *S. purpuratus* (A0A7M7T4C2) (**Figure 6A**). Human CAV1, a member of the Cav1/Cav3 family (Type II-CAV), is shown for comparison (**Figure 6A**). For clarity, dimers rather than higher-order homo-oligomer of each representative caveolin are shown. Our initial analysis was performed using AF2.1, and key findings were examined further using AF2.2.

The Choa-CAV from *S. rosetta* is predicted by AF2.1 to contain a long helix similar in length to its metazoan counterparts with high confidence. However, it is predicted to exist in an extended rather than bow-like shape (**Figure 6B**). It lacks a signature motif, and the hook structure was predicted only with low confidence. Interestingly, AF2.1 predicts multiple copies of Choa-CAV assemble into a parallel helical bundle, and C-terminal β-strands were observed only for 5-mers and higher oligomers (**Figure 6B-D**, **Supplemental file 1**). The structure of Choa-CAV monomer predicted by AF2.2 was generally similar to that suggested by AF2.1. In contrast to the predictions of AF2.1, however, AF2.2 suggests Choa-CAV assembles into spirally packed hat-like structures (**Figure S6**). This model predicts that the protomer undergoes a substantial conformational change as part of the oligomerization process. When we rendered the AF2.2-predicted 7-mer with hydrophobicity, an amphipathic structure emerged reminiscent of the amphipathic feature of the human CAV1 complex (**Figure S6B-D**). Notably, the structure of the oligomeric forms of Choa-CAV were predicted with lower confidence than for the protomer either by AF2.1 or AF2.2.

All caveolins from the representative sponge species *A. queenslandica* were clustered as a clade separated from all other metazoan caveolins (**Table 1**). Intriguingly, caveolins from this clade are predicted by AF2.1 to share features of both Choa-CAV and typical caveolins. All sequences from this clade were predicted to form long bow-like helices (**Table 1, Supplemental file 1**). However, as illustrated in the case of *A. queenslandica* caveolin A0A1X7UHP5, AF2.1 predicts two totally different oligomerization patterns. Some models were very similar in overall appearance to Choa-CAV, whereas others displayed spiral packing features similar to human CAV1 (**Figure 6E-J**). However, the structure of the signature motif was only predicted with low confidence. In addition, no C-terminal β-sheet or closed structures were predicted to form for any *n*-mer we tested for the sequences examined from this clade (**Figure 6H and J, Supplemental file 1**). Here again, the predictions of AF2.2 and AF2.1 were similar for monomers, but diverged for higher order oligomers. As for the case of Choa-CAV, AF2.1 predicts the atypical caveolin A0A1X7UHP5 generates helical bundles (**Figure S7A**), while

AF2.2 predicts higher order oligomers of this and other atypical caveolins range in shape from closed rings to closed discs (**Figure S7B-P**). For the most ring-like structure (A0A1X7TMH4), the entire spoke region protruded perpendicularly to the plane of the membrane bilayer (**Figure S7H-J**). Confidence levels for the predicted oligomeric structures were much lower than for a single protomer. These atypical caveolins may thus also have the ability to assemble into oligomers with variable conformations.

A previous study defined a "CAV-like" clade of caveolins [5]. This clade was separated from all other typical-CAV1 on the updated phylogenetic tree (**Table 1**). Analysis of the predicted structures of the CAV-like caveolins revealed they share some features with other Typical-CAVs, but also exhibit distinct properties (**Table 1**). Most of the analyzed sequences from this clade were predicted to form homo-oligomers with spiral helical regions. Most of them even generated closed disc structures and formed C-terminal β-sheets for trimers and larger oligomers (**Table 1, supplemental file 1, Figure 6K-M and Q**). However, the pin motif was absent from all predicted structures, and no C-terminal β-strands were predicted to be present in dimers (**Table 1, supplemental file 1, Figure 6K-M and Q**). We thus re-named this clade as Type I-CAV and all other Typical-CAVs as Type II-CAV.

Although disc-shaped structures dominated the predictions for Typical-CAVs (**Supplemental file 1**), AF2.2 predicted ring-like structures for two caveolins from typical-CAV clades (**Figure S8A-B**). These structures are very similar to the predicted structures for truncated versions of human CAV1, CAV1 ΔC (residues 1-147) and CAV1 ΔN ΔC (residues 49-147) (**Figure S8C, D**). Importantly, these two truncated versions of CAV1 are capable of driving the formation of caveolae in mammalian cells [69]. This implies that a ring-like amphipathic structure may be sufficient to drive the deformation of membranes and caveolae assembly.

Structural basis for hetero-oligomerization of evolutionarily distinct caveolins

Finally, we examined the ability of evolutionarily distant caveolins to hetero-oligomerize. The best studied example is mammalian CAV1 and CAV2. CAV1 and CAV2 are co-expressed and co-localize in caveolae in many cell types, co-fractionate during purification, and co-migrate in native electrophoresis gel [5,34,38,70,71]. However, in the absence of CAV1, CAV2 fails to form stable high molecular weight oligomers, is localized in the Golgi complex, and cannot support caveolae biogenesis [34,63-65]. Consistent with this, as shown above, CAV2 is predicted to be unable to form regular closed disc-shaped complexes. CAV2 is separated by a long evolutionary distance from CAV1 [5]. It is thus not clear how these two evolutionarily distinct caveolins interact with each other.

To address this question, as a test case we asked AF2 to predict the structure of a complex containing five CAV1 protomers and five CAV2 protomers. AF2 predicted CAV1 and CAV2 indeed can assemble into a hetero-oligomer. Furthermore, the overall structure and organization of the CAV1/CAV2 complex was generally similar to that of the experimentally determined CAV1 8S complex (**Figure 7A-B**). Of the five models AF2 predicted, 3 assembled into fully closed disc shapes with a flat membrane facing surface, and all of them segregated CAV1 and CAV2 protomers on opposite sides of the complex (**Figure S9A-J**). The structure of CAV2 protomers within the complex was predicted with high confidence (**Figure 7F-J**). These findings suggest that CAV1 helps to stabilize CAV2 into a conformation that is compatible with interacting with CAV1 to create a heter-oligomeric complex capable of packing into caveolae. They also raise the interesting possibility that depending on the stoichiometry hetero-oligomeric

caveolin complexes could be asymmetrical structures, with CAV1 and CAV2 endowing distinct functions to different sides of the complexes. It is important to note that the predicted CAV1:CAV2 oligomerization patter represents one of several possible ways these two caveolins could pack within complexes, even when the same number of CAV2 protomers are present in the complex. The organization of complexes could also differ as a function of the ratio of CAV1:CAV2 or for complexes containing different numbers of protomers.

Another notable feature of the CAV1/CAV2 hetero-oligomers is the nature of the interactions between CAV1 and CAV2 at the protomer-protomer interfaces. Due to the spiral packing of protomers into a disc, neighboring CAV1 and CAV2 protomers can potentially interact with one another in two different orientations: CAV1/CAV2 or CAV2/CAV1 (**Figure 7C**). When we overlayed these two different configurations, their scaffolds appeared almost identical except for small differences in angles in the C-terminal region and the N-terminal coiled region (**Figure 7D-E**). Close examination of the pin motif region reveals the exact same set of critical residues that mediate the protomer-protomer interactions in the experimental CAV1 homo-oligomeric structure also exist for both possible arrangements of the CAV1/CAV2 protomer interface. The angles of these residues are also nearly identical to each other and are also very similar to those observed in the CAV1 homo-oligomeric complex model (**Figure 7E**) [56].

Even more elaborate hetero-oligomers may potentially form in cells where CAV1, CAV2, and CAV3 are co-expressed, such as smooth muscle cells. In this cell type, CAV1, CAV2 and CAV3 co-immunoprecipitate [72], raising the interesting possibility that they form mixed hetero-oligomers containing all three caveolin family members. Consistent with this possibility, AF2 predicts that these three caveolins likewise may assemble into a common disc (**Figure S9 and S10**).

Discussion

Although caveolins are expressed in a wide variety of organisms throughout metazoa, very little is known about the structural features of caveolins beyond human CAV1. Based on the proven ability of AF2 to predict overall features of the secondary, tertiary, and quaternary structures of human CAV1 [56], we used it here as a tool to gain insights into the shared and distinct features of caveolins throughout the gene family.

Remarkably, most metazoan caveolins share the ability to assemble into helically-based spiral multimers with a hydrophobic membrane facing surface. Yet, many species, including humans, express multiple, evolutionarily divergent forms of the protein (**Table 1**). Presumably, these different caveolins fulfill tissue- or cell type-specific roles in distinct organisms. Mammalian CAV3, a muscle-specific form of the protein, is a good example. Both mammalian CAV1 and CAV3 are capable of forming caveolae, albeit with the help of different members of the cavin family [45,73,74]. Like CAV1, CAV3 has been reported to form disc-shaped complexes [33]. These complexes are similar but not identical to those formed by CAV1 experimentally [61] and as predicted here by AF2. Exactly how the differences in the structures of CAV1 versus CAV3 dictate their tissue-specific functions, however, remains to be determined.

It is also important to note that not all of the caveolins we examined are predicted to form closed discs. This emphasizes that caveolins do not necessarily need to form intact homo-oligomeric discs to function. This is well illustrated by the case of CAV2, which gives rise to severe phenotypes when knocked out in mice [40]. One simple mechanism by which such caveolins function is through hetero-oligomerization. Hetero-oligomerization could endow unique properties to caveolins complexes, as predicted here for the case of asymmetric CAV1/CAV2

complexes. The exact organization of such complexes likely vary depending on the relative abundance of each caveolin family member. It is also possible that other non-disc forming caveolins may have distinct functions as monomers or homo-oligomers. These possibilities will require further investigation.

Classically, caveolins have been depicted as consisting of several major domains including a scaffolding domain, an oligomerization domain, an intramembrane domain, and a C-terminal region. Based on our structure-based analysis, we now propose that caveolins can be decomposed into six distinct structural units: a variable N-terminal region, pin motif, hook structure, bow-like helical region, β-strand, and variable C-terminal region. It is interesting to compare the key similarities and differences in these features across major caveolin groups (**Figure 6**). The hook structure and long helices appear to be the most highly conserved structural features across most caveolins. The helical region appears to be especially critical, as caveolins that lack helices are predicted to oligomerize poorly (**Table 1**). The pin motif also likely helps stabilize the complexes to generate closed disc-like structures but may not be essential for certain types of caveolins. Based on sequence alignments and AF2 predictions, the structure and packing of the pin motif appear to be conserved among many metazoan caveolins but are missing from caveolins from *A. queenslandica* (**Table 1**, **Table S1**, **Supplemental file 1**). These caveolins are predicted by AF2.1 to fail to pack into spirally assembled complexes containing more than four monomers (**Supplemental file 1**). However, four caveolins from *S. purpuratus* also lack well defined pin motifs, yet are strongly predicted by AF2 to form disc-like complexes (**Table 1**). This suggests different structural elements of caveolins could potentially compensate for one another during evolution to support complex formation. Assembly of the C-terminal β-strands into a central parallel β-barrel further contributes to the proper packing of protomers into regularly shaped disc-shaped complexes [61]. Absence or disruption of the β-barrel destabilizes the complexes and can interfere with caveolin function, as evidenced by C-terminal frameshift mutations of CAV1 linked to pulmonary arterial hypertension and congenital generalized lipodystrophy [53,70,71,75-79]. This modular design could also help to support hetero-oligomer formation. Considering the large number of possible combinations among distinct caveolins in species expressing multiple forms of the protein, this could potentially be an important mechanism to fulfill different functions during development or in response to varying environmental cues.

Many caveolins are predicted to contain N- and C-terminal unstructured regions, and these regions can vary dramatically in length and composition (**Table 1**). In many caveolins, the N-terminal regions appear to be longer and more variable than the C-terminal regions. Choa-CAVs and Atypical-CAVs tend to have longer C-terminal regions, whereas type II-CAV exhibit considerable diversity in this region (**Table 1**). In human CAV1 the C-terminal variable region is essentially absent, but in other caveolins, this region is predicted to be upwards of 45 residues in length. The N-terminal region of CAV1 differs in length between the two major isoforms, CAV1α and CAV1β [80,81]. It also differs between CAV1 and CAV3, which we and others report to contain an N-terminal α-helix [62]. While residues 1-48 of mammalian CAV1 are dispensable for caveolae formation [5], they fulfill several regulatory roles. Tyr 14, a known Src phosphorylation site in CAV1, is required for development in zebrafish [82]. The N-terminus of CAV1 also contains an additional phosphorylation site and multiple ubiquitination sites and participates in several protein-protein interactions [83-90]. Other potential roles of the disordered regions include aiding in curvature generation and mediating protein-protein interactions via phase separation/ condensate formation [91,92]. There is some evidence that such interactions occur between

disordered regions in Cav1 and cavin-1 [93]. Whether this represents a more general mechanism responsible for mediating other protein-protein interactions with the caveolins will be an important possibility to test moving forward.

Another important finding reported here is that caveolin-related proteins are not limited to Metazoa: they are also found in choanoflagellates, the closest living relatives of animals. Thus, caveolins appear to be among the ancient proteins borrowed and repurposed from unicellular organisms. This raises 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 [68,94,95].

A related question is which structural features of caveolins are optimized for caveolae formation and which fulfill broader functional roles. Current evidence suggests caveolae are limited to vertebrates due to the requirement for both cavelins and cavins to build caveolae [48,96]. This means that outside of vertebrates, caveolins function independently of classically defined caveolae. Even in vertebrates, there are cell types such as neurons where caveolins are expressed, but caveolae are not present due to the absence of cavin expression [44,47]. Some invertebrate caveolins, such as honeybee caveolin, are capable of supporting caveolae formation, whereas others are not [5]. Several of the functions ascribed to caveolins, such as regulation of signaling and lipid metabolism, may not require the incorporation of caveolins into caveolae [44,47]. The mechanosensing and mechanoprotection roles of caveolins [11,12] may also not be limited to organisms where caveolae are present. Recent evidence suggests that caveolins can protect and maintain tissue integrity in the absence of cavins and traditional caveolae in invertebrates [51]. It is also formally possible that expression of caveolins may be sufficient to drive the formation of caveolae-like structures in in other organisms, as has been observed upon ectopic expression of caveolins in *E. coli* [69,97]. Our analysis does, however, suggest that the most evolutionary distant caveolins, such as the Choa caveolins or atypical caveolins found in *A. queenslandica*, are structurally distinct from other caveolins, implying they also function differently from other caveolins. As experimental approaches to study the cell biology of organisms like choanoflagellates continue to advance [98], it should be possible to test the structure-function relationship of these evolutionarily diverse forms of the protein in the future.

Finally, our comparison of caveolins across evolution, especially with the identification of Choa caveolins, predicts a possible path for the evolution of their structure. We speculate ancestral caveolins may have been primarily helical, with a tendency to oligomerize in parallel helical bundles. Early caveolins also likely contained an N-terminal hook-like structure. Incorporation of kinks at key positions may then have allowed the helices to progressively shift to a spiral, partially flattened ring or disc. AF2 hints that this may be the case for the most evolutionarily distant caveolins, although such structures are predicted with low confidence. Our finding that Choa caveolins and Atypical-CAVs can also form ring-like oligomers with a hydrophobic surface suggests this may represent one of the most ancient functional unit of the protein in metazoans. A similar ring structure may also be sufficient to promote membrane bending and caveolae assembly, as evidenced by the predicted structures of human CAV1 lacking N- and C-terminal domains. These possibilities will require further analysis in the future.

There are several limitations to our findings. The most important limitation is that the structure predictions presented here are designed to formulate hypotheses and remain to be verified experimentally. This is especially important because caveolins are a challenging class of proteins for AF2. They fall into the poorly studied class of monotopic membrane proteins, are unrelated in sequence and structure to other proteins, and assemble into symmetric, circular oligomers. Furthermore, the experimental structure of human CAV1 was not present in the database when AF2 was trained. Despite this, AF2 was able to predict the overall fold and many details of the structure of human CAV1 [56], providing the impetus for our current study. Here, we found that both AF2.1 and AF2.2 can sense the differences among different caveolins clades and distinguish them structurally. The exact readouts of AF2.1 and AF2.2 did, however, vary. AF2.1 underestimates the numbers of protomers within the spirally packed human CAV1 oligomers [56], whereas AF2.2 has a stronger tendency than AF2.1 to assemble caveolins into closed structures. Their predictions also differed substantially for the most evolutionarily distant caveolins, although in these cases, the confidence of the predictions was also low. The exact oligomeric state of the caveolins is also uncertain. For the case of CAV1, AF2 predicts additional oligomeric states exist in addition to the experimentally observed 11-mer. In support of this possibility, there is experimental evidence that suggests varied oligomeric states may assemble in cells [28-32]. However, some of these oligomeric states predicted by AF2 appear less likely to exist due to energetic strains [56]. The structures presented here should thus not be interpreted as ground truth but rather as providing a window into the principles of design underlying caveolin structure and function. Nevertheless, our finding that the overall features of the oligomers and complexes are common across caveolins suggests that these models may depict their fundamental structural properties.

Methods

AlphaFold2 Prediction

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 [55].

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 [99] was used for alignment image typesetting and exporting.

HMMER searches

For HMMER searches, a hidden Markov model (HMM) profile was built directly from the alignment of truncated caveolin sequences by Kirkham et al 2008 using hmmbuild from the HMMER package [100]. The profile was then searched against the *S. rosetta* and *A. queenslandica* genomes using hmmsearch.

Phylogenetic analyses

The selected caveolin homologues were aligned with MAFFT v7.310 [101,102]. 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) [5] and re-aligning with MAFFT. ProtTest3 (version 3.4.2) [103] was then used to determine the best model of evolution (LG+I+Γ+F). Finally, a maximum likelihood phylogeny was inferred using RAxML (version 8.2.11; [104,105] with 100 rapid bootstraps. A trimmed tree was produced in RStudio using the *ape* package [106]. Trees were rendered using FigTree v1.4.4 and prepared for publication using Inkscape (version 1.0.1).

Acknowledgments

We thank Mr. Fuchang Han from Kaifeng City for the schematic diagrams of different species for Figure 3, Dr. Hassane Mchaourab for helpful discussions, and Dr. Melanie Ohi for critical feedback on an early version of the manuscript. This work was supported by National Institutes of Health grant R01 HL144131 (AKK), National Institutes of Health grant R01GM080403 (JM), National Institutes of Health grant R01HL122010 (JM), National Institutes of Health grant R01GM129261 (JM), Humboldt Professorship of the Alexander von Humboldt Foundation (JM), and HHMI (to Jochen Zimmer). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

Conceptualization: BH, AKK
Methodology: BH, AG, LFLW
Investigation: BH, LFLW
Data curation: BH
Formal analysis: BH, LFLW
Validation: AG
Writing—Original draft: BH, AKK, LFLW
Writing—Review and editing: BH, LFLW, AG, JM, EK, AKK
Visualization: BH, LFLW
Supervision: AKK, JM, EK

Project administration: AKK
Funding acquisition: AKK, JM

Declaration of interests. The authors declare no competing interests.

Figure Legends

Figure 1. Predicted structures of human CAV1, CAV2, and CAV3 monomers compared to the experimental structure of human CAV1. (A) Sequence alignment of human CAV1, CAV2, and CAV3. Specific regions of each are color coded as follows: pin motif, yellow; signature motif, red; scaffolding domain, green; intramembrane domain, purple; spoke-like region, gray; β-strand, cyan. The oligomerization domain includes the signature motif and scaffolding domain and is indicated by brackets. **(B-C)** *En face* and side views of the cryo-EM structure of the human 8S CAV1 complex (PDB: 7SC0). **(D-I)** Predicted structures of CAV1 (D, E), CAV2 (F, G) and CAV3 (H, I) monomers from the AF2 database. Structures in C, E, G, and I are rotated 90˚ on y-axis relative to those shown in B, D, F, and H. Color coding in (B-I) are is identical to that in panel A. **(J-K)** As in B and C except that the position of a single protomer is highlighted in red. **(L-Q)** As in D-I, except the structures are colored by pLDDT confidence values.

Figure 2. Predicted structures of human CAV1, CAV2, and CAV3 oligomers. AF2.2 models for 11-mers of human CAV1 (**A-C**), CAV2 **(D-F),** and CAV3 **(G-I).** Views of the cytoplasmic surface, side view, and membrane facing surface are shown for each. Models are colored according to the pLDDT confidence values. Zoomed in views of the C-terminal regions of each model are highlighted in the lower panels in B, E, and H.

Figure 3. Conserved structural features of metazoan caveolins highlighted by AF2 predictions. AF2.2 models for monomers **(A)** and multimers **(B-E)** of metazoan caveolins from representative species of thirteen different metazoan phyla/superphyla. Models in rows A and B are colored by pLDDT confidence values. Models in rows C-E are colored by lipophilicity values. All multimer models are for 11-mers except A0A1X7UHP5 which is 10-mer.

Figure 4. Proposed structural elements of metazoan caveolins. Structural elements include a variable N-terminal region (red), pin motif (yellow), hook structure (green), bow-like helical region (blue), β-strand (orange), and variable C-terminal 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 8S CAV1 complex (PDB: 7SC0), **(B)** a 2-mer of human CAV1 (Q03135) predicted by AF2.1, and **(C)** a 2-mer of *Caenorhabditis elegans* caveolin (Q94051) predicted by AF2.1.

Figure 5. Phylogenetic relationships of caveolin sequences. A maximum likelihood phylogeny was inferred for representative caveolins from the current study (black) in combination with caveolin sequences previously analyzed by [5]. Previously analyzed caveolins are color coded according to their classifications [5]. Support values (percentage replication) are shown for the major splits. *, caveolin-related protein from *Salpingoeca rosetta*.

Figure 6. Structural features of representative metazoan and non-metazoan caveolins. (A) Sequence alignment of representative caveolins. Conserved residues are colored according to the percent identity. **(B-P)** Panoramic and zoomed in views of the N- and C-terminal regions of AF2.1 models of dimers from representative caveolins**. (B-D)** Choa-CAV from *Salpingoeca rosetta* (F2U793). **(E-J)** Rank 1 and rank 3 models of an Atypical-CAV from *Amphimedon*

queenslandica (A0A1X7UHP5). Residue 1-160 were hidden in (F), (G), (I) and (J) to better display the structures. **(K-M)** Type I-CAV from *Strongylocentrotus purpuratus* (A0A7M7T4C2). **(N-P)** Type II-CAV from human CAV1. In B to P, all models are colored by pLDDT confidence values. Pin motif in (O) is highlighted by gray dotted lines. Regions in (I) and (L) which align with the pin motif of human CAV1 are highlighted by gray dotted lines. Side chains within signature motif region and the hydrogen bonds associated with them are shown in (C), (F), (I), (L) and (O). **(Q)** 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 1.

Figure 7. Predicted structures of hetero-oligomeric CAV1/CAV2 complexes. **(A-B)** *En face* and tilted views of a hetero-oligomeric caveolin complex containing five human CAV1 protomers (purple) and five human CAV2 protomers (orange) from a secondary structure model predicted by AF2.2. **(C)** CAV1 and CAV2 interact with each other in two different ways within the complex, as highlighted here by extracting them from the full complex in B. **(D)** Overlay of the CAV1/CAV2 heterodimers from C. (**E**) Zoomed in view of the boxed region in D. **(F-J)** As in A-J, except the models are colored by pLDDT values.

Table 1. Summary of the structural features of caveolins suggested by AF2.1 predictions. Phylogenetic tree shown on the left-hand side of the table is based on Figure 5. Since not all caveolins were predicted to form closed disc-like structures, the presence or absence of structural features was scored from predicted homo-dimers. When a structural element was not predicted in a dimer, predictions for higher order n-mers were checked and used as supplementary information in this table. ●, structural element was predicted to be present in a dimer; ◑, structural element was not predicted in dimers but was predicted for other *n*-mers; ○, structural structure was not predicted for any *n*-mers. NA, not applicable. In the C-terminal beta sheet column (C-term sheet), the numbers in parentheses indicate how many residues were predicted to form the beta strand. The numbers in N-terminal variable regions (N-term VR) and C-terminal variable regions (C-term VR) columns represent the number or residues in this region based on alignments with the sequence of human CAV1.

Scoring rules of the “Spiral assembly” were as follows: ☆, long helices lack a bow-like feature even at monomer level and cannot be assembled as homo-oligomers into a spiral packing pattern; ★, long helices are predicted to be bow-like in monomers, but cannot be assembled as homo-oligomers into a spiral packing pattern; ★★, long helices are predicted to be bow-like and can assemble as homo-oligomers into a loose spiral packing pattern, but do not form closed disc-like structures for any n-mer; ★★★, protomers homo-oligomerize into closely packed spirals, but do not form closed discs for any n-mer; ★★★½, closed disc-like structures are predicted to form, but the predictions were based on the truncated version of corresponding caveolin and the disc-like models were not the rank 1 model; ★★★★, closed disc-like structures were predicted, but none of the closed models were the rank 1 model; ★★★★½, closed disc-like structures were predicted as rank 1 models, but the predictions were based on truncated version of corresponding caveolin; ★★★★★, closed disc-like structures were predicted as rank 1 models, and the predictions were based on the full-length sequence of the corresponding caveolin.

In the “NC (negative charge) sites on bottom” column, the superscript ^ indicates that multiple negative charge sites from one protomer were found on the bilayer facing side and # indicates that positive charge sites were found on the bilayer facing side.

Supplementary materials

Figure S1. Structures of human CAV1, CAV2, and CAV3 oligomers predicted by AF2.1 and AF2.2. (A-I) 8-mers models of CAV1 (A-C), CAV2 (D-F) and CAV3 (G-I) predicted by AF2.1. **(J-R)** 11-mer models of CAV1 (J-L), CAV2 (M-O) and CAV3 (P-R) oligomers predicted by AF2.2. All the models are colored with the corresponding pLDDT confidence values. Zoomed in views of the C-terminus regions of each model are highlighted in the lower panels in B, E, H, K, N and Q.

Figure S2. Prediction results for the structure of human CAV2 over time. Results are shown for both AF2.1 and AF2.2 predictions performed during the indicated timeframes. All models were colored by pLDDT confidence values.

Figure S3. Conserved structural features of metazoan caveolins highlighted by AF2.1 predictions. AF2.1 models for monomers **(A)** and multimers **(B-E)** of metazoan caveolins from representative species of thirteen different metazoan phyla/superphyla. Models in rows A and B are colored by pLDDT confidence values. Models in rows C-E are colored by lipophilicity values. Sizes of oligomers vary between 4-mers and 9-mers.

Figure S4. Electrostatic potential distribution patterns on the bilayer facing surface of predicted caveolin oligomers. Examples of different patterns of charged residues on the membrane facing surface are shown for representative caveolins oligomers predicted by AF2.2. They include **(A)** completely neutral membrane facing surface; **(B-C)** a negatively charged ring contributed by a single Glu located in the middle of the long helices; **(D)** a negatively charged ring contributed by a single Glu located near the C-terminal region of the long helices; **(E)** two negatively charged rings contributed by a Glu or an Asp in the middle of the long helices; **(F)** a single negatively charged ring contributed by a Glu and an Asp in the middle of the long helices, **(G)** a positively charged ring contributed by a single Lys in the middle of the long helices, **(H)** a positively charged ring contributed by a single Lys near the C-terminal region of the long helices. The relative abundance of each pattern for is listed below each model (from a total of 72 caveolins investigated).

Figure S5. Structural relationship between the hook structure and signature motif in the cryo-EM structure of the human CAV1 8S complex. (A) Two adjacent protomers within the complex are highlighted in green and blue. **(B)** Zoomed in view of the hook structure (pink) and signature motif (purple) from (A).

Figure S6. Structures of Choa-CAV (F2U793) predicted by AF2.1 and AF2.2. **(A)** Comparison of the prediction results for the indicated *n*-mers for AF2.1 versus AF2.2. **(B-D)** Three different views of Choa-CAV 7-mer predicted by AF2.2. Models in panel A are colored by pLDDT values and models in panels B-D are colored by lipophilicity values.

Figure S7. Structures of Atypical-CAVs predicted by AF2.1 and AF2.2. **(A)** Comparison of predicted structures for the indicated *n*-mers for AF2.1 versus AF 2.2 for a representative Atypical-CAV (A0A1X7UHP5). The sequence of the N-terminal variable region was removed to avoid going over the sequence input limit for the AF2.1-based predictions of 5- to 7-mers. **(B-P)** Predicted structures of several different atypical caveolins, as modeled by AF2.2. Three different views are shown for each. **(B-D)** A0A1X7UHP5 10-mer, **(E-G)** A0A1X7UGA1 11-mer;

(H-J) A0A1X7TMH4 11-mer; **(K-M)** A0A1X7VPY7 11-mer; and **(N-P)** A0A1X7VRV8 11-mer. All models were colored by pLDDT values.

Figure S8. Examples of ring-like caveolin oligomers predicted by AF2.2. **(A)** 11-mer of a Type I-CAV (R7U639). **(B)** 11-mer of a Type II-CAV (A0A267DC90). (**C**) 11-mer models of human CAV1 ΔC (Residue 1-147) oligomer. **(D)** 11-mer models of human CAV1 ΔNΔC (residues 49-147) oligomer.

Figure S9. Top five predicted structures of hetero-oligomeric complexes formed by human CAV1/CAV2 or CAV1/CAV2/CAV3. (A-J) *En face* and side views of human CAV1/CAV2 hetero-oligomeric complexes formed by 5 CAV1s (purple) and 5 CAV2s (orange) as predicted by AF2.2. All 5 of the top ranked models are shown. **(K-T)** *En face* and side views of human CAV1/CAV2/CAV3 hetero-oligomeric complexes formed by 3 CAV1s (purple), 3 CAV2s (orange), and 3 CAV3s (gray) as predicted by AF2.2. All 5 of the top ranked models are shown. Models in upper panel are colored by pLDDT confidence value, and models in lower panels are colored by different caveolins.

Figure S10. Predicted structures of hetero-oligomeric CAV1/CAV2/CAV3 complexes. **(A-B)** *En face* and tilted views of a hetero-oligomeric caveolin complex containing three human CAV1 protomers (purple), three human CAV2 protomers (orange), and three human CAV3 protomers (gray) from a secondary structure model predicted by AF2.2. **(C)** CAV1, CAV2, and CAV3 interact with each other in several different ways within the complex, as highlighted here by extracting them from the full complex in B. **(D)** Overlay of the heterodimers shown in panel C. (**E**) Zoomed in view of the boxed region in D. **(F-J)** As in A-J, except the models are colored by pLDDT values.

Table S1. Alignment of the caveolin sequences clustered in Figure 5. Sequences were truncated to a region corresponding to residues 54–158 in human CAV1 prior to alignment. Residues were colored by percentage identity. Four standard annotations for the alignment (conservation, quality, occupancy and consensus) are shown below the sequences.

File S1. Prediction results of the structures of caveolins monomer or oligomers by AlphaFold2.1 with different protomer numbers. The 2D example color sketches were generated from the 3D model by AlphaFold2_advanced notebook. All right panels of the 2D schematics are colored by pLDDT confidence values. For the left panels, the monomers were colored by N-term to C-term, the multimers were colored by chain. Five models were generated for each prediction. The 2D sketches were based on rank 1 models (R1) in this supplemental file if there is no special note was left under the pLDDT value.

References

1. Rout, M.P., and Field, M.C. (2017). The Evolution of Organellar Coat Complexes and Organization of the Eukaryotic Cell. Annu Rev Biochem *86*, 637-657. 10.1146/annurev-biochem-061516-044643.

2. Dacks, J.B., and Robinson, M.S. (2017). Outerwear through the ages: evolutionary cell biology of vesicle coats. Curr Opin Cell Biol *47*, 108-116. 10.1016/j.ceb.2017.04.001.

3. Attwood, M.M., Krishnan, A., Almen, M.S., and Schioth, H.B. (2017). Highly diversified expansions shaped the evolution of membrane bound proteins in metazoans. Sci Rep *7*, 12387. 10.1038/s41598-017-11543-z.

4. Field, M.C., Gabernet-Castello, C., and Dacks, J.B. (2007). Reconstructing the evolution of the endocytic system: insights from genomics and molecular cell biology. Adv Exp Med Biol *607*, 84-96. 10.1007/978-0-387-74021-8_7.

5. Kirkham, M., Nixon, S.J., Howes, M.T., Abi-Rached, L., Wakeham, D.E., Hanzal-Bayer, M., Ferguson, C., Hill, M.M., Fernandez-Rojo, M., Brown, D.A., et al. (2008). Evolutionary analysis and molecular dissection of caveola biogenesis. J Cell Sci *121*, 2075-2086. jcs.024588 [pii]

10.1242/jcs.024588.

6. Spisni, E., Tomasi, V., Cestaro, A., and Tosatto, S.C. (2005). Structural insights into the function of human caveolin 1. Biochem Biophys Res Commun *338*, 1383-1390. S0006-291X(05)02349-1 [pii]

10.1016/j.bbrc.2005.10.099.

7. Williams, T.M., and Lisanti, M.P. (2004). The Caveolin genes: from cell biology to medicine. Ann Med *36*, 584-595. 10.1080/07853890410018899.

8. Bastiani, M., and Parton, R.G. (2010). Caveolae at a glance. J Cell Sci *123*, 3831-3836. 123/22/3831 [pii]

10.1242/jcs.070102.

9. Parton, R.G. (2018). Caveolae: structure, function, and relationship to disease. Annu Rev Cell Dev Biol *34*, 111-136. 10.1146/annurev-cellbio-100617-062737.

10. Chidlow, J.H., Jr., and Sessa, W.C. (2010). Caveolae, caveolins, and cavins: complex control of cellular signalling and inflammation. Cardiovasc Res *86*, 219-225. cvq075 [pii]

10.1093/cvr/cvq075.

11. Del Pozo, M.A., Lolo, F.N., and Echarri, A. (2020). Caveolae: Mechanosensing and mechanotransduction devices linking membrane trafficking to mechanoadaptation. Curr Opin Cell Biol *68*, 113-123. 10.1016/j.ceb.2020.10.008.

12. Parton, R.G., and del Pozo, M.A. (2013). Caveolae as plasma membrane sensors, protectors and organizers. Nat Rev Mol Cell Biol *14*, 98-112. nrm3512 [pii]

10.1038/nrm3512.

13. Nassoy, P., and Lamaze, C. (2012). Stressing caveolae new role in cell mechanics. Trends Cell Biol *22*, 381-389. S0962-8924(12)00074-8 [pii]

10.1016/j.tcb.2012.04.007.

14. Lamaze, C., Tardif, N., Dewulf, M., Vassilopoulos, S., and Blouin, C.M. (2017). The caveolae dress code: structure and signaling. Curr Opin Cell Biol *47*, 117-125. 10.1016/j.ceb.2017.02.014.

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

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

17. Root, K.T., Plucinsky, S.M., and Glover, K.J. (2015). Recent progress in the topology, structure, and oligomerization of caveolin: a building block of caveolae. Curr Top Membr *75*, 305-336. S1063-5823(15)00008-3 [pii]

10.1016/bs.ctm.2015.03.007.

18. Mercier, I., Jasmin, J.F., Pavlides, S., Minetti, C., Flomenberg, N., Pestell, R.G., Frank, P.G., Sotgia, F., and Lisanti, M.P. (2009). Clinical and translational implications of the caveolin gene family: lessons from mouse models and human genetic disorders. Lab Invest *89*, 614-623.

19. Drab, M., Verkade, P., Elger, M., Kasper, M., Lohn, M., Lauterbach, B., Menne, J., Lindschau, C., Mende, F., Luft, F.C., et al. (2001). Loss of caveolae, vascular dysfunction, and pulmonary defects in caveolin-1 gene-disrupted mice. Science *293*, 2449-2452.

20. Razani, B., Engelman, J.A., Wang, X.B., Schubert, W., Zhang, X.L., Marks, C.B., Macaluso, F., Russell, R.G., Li, M., Pestell, R.G., et al. (2001). Caveolin-1 null mice are viable but show evidence of hyperproliferative and vascular abnormalities. J Biol Chem *276*, 38121-38138.

21. Fra, A.M., Williamson, E., Simons, K., and Parton, R.G. (1995). De novo formation of caveolae in lymphocytes by expression of VIP21-caveolin. Proc. Natl. Acad. Sci. USA *92*, 8655-8659.

22. Zhao, Y.Y., Liu, Y., Stan, R.V., Fan, L., Gu, Y., Dalton, N., Chu, P.H., Peterson, K., Ross, J., Jr., and Chien, K.R. (2002). Defects in caveolin-1 cause dilated cardiomyopathy and pulmonary hypertension in knockout mice. Proc Natl Acad Sci U S A *99*, 11375-11380. 10.1073/pnas.172360799

172360799 [pii].

23. Williams, T.M., and Lisanti, M.P. (2004). The caveolin proteins. Genome Biol *5*, 214. 10.1186/gb-2004-5-3-214.

24. Busija, A.R., Patel, H.H., and Insel, P.A. (2017). Caveolins and cavins in the trafficking, maturation, and degradation of caveolae: implications for cell physiology. Am J Physiol Cell Physiol *312*, C459-C477. ajpcell.00355.2016 [pii]

10.1152/ajpcell.00355.2016.

25. Way, M., and Parton, R.G. (1995). M-Caveolin, a Muscle-Specific Caveolin-Related Protein. FEBS Letters *376*, 108-112.

26. Tang, Z., Scherer, P.E., Okamoto, T., Song, K., Chu, C., Kohtz, D.S., Nishimoto, I., Lodish, H.F., and Lisanti, M.P. (1996). Molecular cloning of caveolin-3, a novel member of the caveolin gene family expressed predominantly in muscle. J Biol Chem *271*, 2255-2261.

27. Galbiati, F., Engelman, J.A., Volonte, D., Zhang, X.L., Minetti, C., Li, M., Hou, H., Jr., Kneitz, B., Edelmann, W., and Lisanti, M.P. (2001). Caveolin-3 null mice show a loss of caveolae, changes in the microdomain distribution of the dystrophin-glycoprotein complex, and t-tubule abnormalities. J Biol Chem *276*, 21425-21433. 10.1074/jbc.M100828200

M100828200 [pii].

28. Fernandez, I., Ying, Y., Albanesi, J., and Anderson, R.G. (2002). Mechanism of caveolin filament assembly. Proc Natl Acad Sci U S A *99*, 11193-11198. 10.1073/pnas.172196599

172196599 [pii].

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

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

31. Khater, I.M., Meng, F., Wong, T.H., Nabi, I.R., and Hamarneh, G. (2018). Super resolution network analysis defines the molecular architecture of caveolae and caveolin-1 scaffolds. Sci Rep *8*, 9009. 10.1038/s41598-018-27216-4.

32. Khater, I.M., Liu, Q., Chou, K.C., Hamarneh, G., and Nabi, I.R. (2019). Super-resolution modularity analysis shows polyhedral caveolin-1 oligomers combine to form scaffolds and caveolae. Sci Rep *9*, 9888. 10.1038/s41598-019-46174-z.

33. Whiteley, G., Collins, R.F., and Kitmitto, A. (2012). Characterization of the molecular architecture of human caveolin-3 and interaction with the skeletal muscle ryanodine receptor. J Biol Chem *287*, 40302-40316. M112.377085 [pii]

10.1074/jbc.M112.377085.

34. Hayer, A., Stoeber, M., Bissig, C., and Helenius, A. (2010). Biogenesis of caveolae: stepwise assembly of large caveolin and cavin complexes. Traffic *11*, 361-382. TRA1023 [pii]

10.1111/j.1600-0854.2009.01023.x.

35. Ren, X., Ostermeyer, A.G., Ramcharan, L.T., Zeng, Y., Lublin, D.M., and Brown, D.A. (2004). Conformational defects slow Golgi exit, block oligomerization, and reduce raft affinity of caveolin-1 mutant proteins. Mol Biol Cell *15*, 4556-4567.

36. Park, D.S., Woodman, S.E., Schubert, W., Cohen, A.W., Frank, P.G., Chandra, M., Shirani, J., Razani, B., Tang, B., Jelicks, L.A., et al. (2002). Caveolin-1/3 double-knockout mice are viable, but lack both muscle and non-muscle caveolae, and develop a severe cardiomyopathic phenotype. Am J Pathol *160*, 2207-2217.

37. Scherer, P.E., Okamoto, T., Chun, M., Nishimoto, I., Lodish, H.F., and Lisanti, M.P. (1996). Identification, sequence, and expression of caveolin-2 defines a caveolin gene family. Proc Natl Acad Sci U S A *93*, 131-135.

38. Scherer, P.E., Lewis, R.Y., Volonte, D., Engelman, J.A., Galbiati, F., Couet, J., Kohtz, D.S., van Donselaar, E., Peters, P., and Lisanti, M.P. (1997). Cell-type and tissue-specific expression of caveolin-2. Caveolins 1 and 2 co-localize and form a stable hetero-oligomeric complex in vivo. J Biol Chem *272*, 29337-29346.

39. de Almeida, C.J.G. (2017). Caveolin-1 and Caveolin-2 Can Be Antagonistic Partners in Inflammation and Beyond. Front Immunol *8*, 1530. 10.3389/fimmu.2017.01530.

40. Razani, B., Wang, X.B., Engelman, J.A., Battista, M., Lagaud, G., Zhang, X.L., Kneitz, B., Hou, H., Jr., Christ, G.J., Edelmann, W., and Lisanti, M.P. (2002). Caveolin-2-deficient mice show evidence of severe pulmonary dysfunction without disruption of caveolae. Mol Cell Biol *22*, 2329-2344.

41. Sowa, G. (2011). Novel insights into the role of caveolin-2 in cell- and tissue-specific signaling and function. Biochem Res Int *2011*, 809259. 10.1155/2011/809259.

42. Parton, R.G., Tillu, V., McMahon, K.A., and Collins, B.M. (2021). Key phases in the formation of caveolae. Curr Opin Cell Biol *71*, 7-14. 10.1016/j.ceb.2021.01.009.

43. Parton, R.G., McMahon, K.A., and Wu, Y. (2020). Caveolae: formation, dynamics, and function. Curr Opin Cell Biol *65*, 8-16. 10.1016/j.ceb.2020.02.001.

44. Pol, A., Morales-Paytuvi, F., Bosch, M., and Parton, R.G. (2020). Non-caveolar caveolins - duties outside the caves. J Cell Sci *133*. 10.1242/jcs.241562.

45. Kovtun, O., Tillu, V.A., Ariotti, N., Parton, R.G., and Collins, B.M. (2015). Cavin family proteins and the assembly of caveolae. J Cell Sci *128*, 1269-1278. 128/7/1269 [pii]

10.1242/jcs.167866.

46. Ariotti, N., and Parton, R.G. (2013). SnapShot: caveolae, caveolins, and cavins. Cell *154*, 704-704 e701. S0092-8674(13)00880-5 [pii]

10.1016/j.cell.2013.07.009.

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

48. Hansen, C.G., and Nichols, B.J. (2010). Exploring the caves: cavins, caveolins and caveolae. Trends Cell Biol *20*, 177-186. S0962-8924(10)00016-4 [pii]

10.1016/j.tcb.2010.01.005.

49. Tang, Z., Okamoto, T., Boontrakulpoontawee, P., Katada, T., Otsuka, A.J., and Lisanti, M.P. (1997). 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. 10.1074/jbc.272.4.2437.

50. Frank, P.G., and Lisanti, M.P. (2006). Zebrafish as a novel model system to study the function of caveolae and caveolin-1 in organismal biology. Am J Pathol *169*, 1910-1912. 10.2353/ajpath.2006.060923.

51. Bhattachan, P., Rae, J., Yu, H., Jung, W., Wei, J., Parton, R.G., and Dong, B. (2020). Ascidian caveolin induces membrane curvature and protects tissue integrity and morphology during embryogenesis. FASEB J *34*, 1345-1361. 10.1096/fj.201901281R.

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

53. Porta, J.C., Han, B., Gulsevin, A., Chung, J.M., Peskova, Y., Connolly, S., McHaourab, H.S., Meiler, J., Karakas, E., Kenworthy, A.K., and Ohi, M.D. (2022). Molecular architecture of the human caveolin-1 complex. Sci Adv *8*, eabn7232. 10.1126/sciadv.abn7232.

54. Ohi, M.D., and Kenworthy, A.K. (2022). Emerging Insights into the Molecular Architecture of Caveolin-1. J Membr Biol. 10.1007/s00232-022-00259-5.

55. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.

56. Gulsevin, A., Han, B., Porta, J.C., Mchaourab, H.S., Meiler, J., and Kenworthy, A.K. (2022). Template-free prediction of a new monotopic membrane protein fold and oligomeric assembly by Alphafold2. Biophys J *in press*. DOI:https://doi.org/10.1016/j.bpj.2022.11.011.

57. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Zidek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590-596. 10.1038/s41586-021-03828-1.

58. Li, S., Okamoto, T., Chun, M., Sargiacomo, M., Casanova, J.E., Hansen, S.H., Nishimoto, I., and Lisanti, M.P. (1995). Evidence for a regulated interaction between heterotrimeric G proteins and caveolin. J Biol Chem *270*, 15693-15701. 10.1074/jbc.270.26.15693.

59. Schlegel, A., Schwab, R.B., Scherer, P.E., and Lisanti, M.P. (1999). 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.

60. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.2010.2004.463034. 10.1101/2021.10.04.463034.

61. Han, B., Porta, J.C., Hanks, J.L., Peskova, Y., Binshtein, E., Dryden, K., Claxton, D.P., McHaourab, H.S., Karakas, E., Ohi, M.D., and Kenworthy, A.K. (2020). Structure and assembly of CAV1 8S complexes revealed by single particle electron microscopy. Sci Adv *6*, eabc6185. 10.1126/sciadv.abc6185.

62. Morales-Paytuvi, F., Ruiz-Mirapeix, C., Fajardo, A., Rae, J., Bosch, M., Enrich, C., Collins, B., Parton, R.G., and Pol, A. (2022). Proteostatic regulation of caveolins avoids premature oligomerisation and preserves ER homeostasis. bioRxiv. bioRxiv 2022.04.24.489297; doi: https://doi.org/10.1101/2022.04.24.489297.

63. Li, S., Galbiati, F., Volonte, D., Sargiacomo, M., Engelman, J.A., Das, K., Scherer, P.E., and Lisanti, M.P. (1998). Mutational analysis of caveolin-induced vesicle formation. Expression of caveolin-1 recruits caveolin-2 to caveolae membranes. FEBS Lett *434*, 127-134. S0014579398009454 [pii].

64. Parolini, I., Sargiacomo, M., Galbiati, F., Rizzo, G., Grignani, F., Engelman, J.A., Okamoto, T., Ikezu, T., Scherer, P.E., Mora, R., et al. (1999). 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.

65. Mora, R., Bonilha, V.L., Marmorstein, A., Scherer, P.E., Brown, D., Lisanti, M.P., and Rodriguez-Boulan, E. (1999). 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.

66. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E., Mitros, T., Richards, G.S., Conaco, C., Dacre, M., Hellsten, U., et al. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. Nature *466*, 720-726. 10.1038/nature09201.

67. Fairclough, S.R., Chen, Z., Kramer, E., Zeng, Q., Young, S., Robertson, H.M., Begovic, E., Richter, D.J., Russ, C., Westbrook, M.J., et al. (2013). Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. Genome Biol *14*, R15. 10.1186/gb-2013-14-2-r15.

68. Shalchian-Tabrizi, K., Minge, M.A., Espelund, M., Orr, R., Ruden, T., Jakobsen, K.S., and Cavalier-Smith, T. (2008). Multigene phylogeny of choanozoa and the origin of animals. PLoS One *3*, e2098. 10.1371/journal.pone.0002098.

69. Ariotti, N., Rae, J., Leneva, N., Ferguson, C., Loo, D., Okano, S., Hill, M.M., Walser, P., Collins, B.M., and Parton, R.G. (2015). Molecular characterization of caveolin-induced membrane curvature. J Biol Chem *290*, 24875-24890. M115.644336 [pii]

10.1074/jbc.M115.644336.

70. Han, B., Copeland, C.A., Kawano, Y., Rosenzweig, E.B., Austin, E.D., Shahmirzadi, L., Tang, S., Raghunathan, K., Chung, W.K., and Kenworthy, A.K. (2016). Characterization

of a caveolin-1 mutation associated with both pulmonary arterial hypertension and congenital generalized lipodystrophy. Traffic *17*, 1297-1312. 10.1111/tra.12452.

71. Copeland, C.A., Han, B., Tiwari, A., Austin, E.D., Loyd, J.E., West, J.D., and Kenworthy, A.K. (2017). 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. 10.1091/mbc.E17-06-0421.

72. Woodman, S.E., Cheung, M.W., Tarr, M., North, A.C., Schubert, W., Lagaud, G., Marks, C.B., Russell, R.G., Hassan, G.S., Factor, S.M., et al. (2004). Urogenital alterations in aged male caveolin-1 knockout mice. J Urol *171*, 950-957. 10.1097/01.ju.0000105102.72295.b8.

73. Bastiani, M., Liu, L., Hill, M.M., Jedrychowski, M.P., Nixon, S.J., Lo, H.P., Abankwa, D., Luetterforst, R., Fernandez-Rojo, M., Breen, M.R., et al. (2009). MURC/Cavin-4 and cavin family members form tissue-specific caveolar complexes. J Cell Biol *185*, 1259-1273. jcb.200903053 [pii]

10.1083/jcb.200903053.

74. Lo, H.P., Nixon, S.J., Hall, T.E., Cowling, B.S., Ferguson, C., Morgan, G.P., Schieber, N.L., Fernandez-Rojo, M.A., Bastiani, M., Floetenmeyer, M., et al. (2015). The caveolin-cavin system plays a conserved and critical role in mechanoprotection of skeletal muscle. J Cell Biol *210*, 833-849. jcb.201501046 [pii]

10.1083/jcb.201501046.

75. Austin, E.D., Ma, L., LeDuc, C., Berman Rosenzweig, E., Borczuk, A., Phillips, J.A., 3rd, Palomero, T., Sumazin, P., Kim, H.R., Talati, M.H., et al. (2012). Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. Circ Cardiovasc Genet *5*, 336-343. CIRCGENETICS.111.961888 [pii]

10.1161/CIRCGENETICS.111.961888.

76. Garg, A., Kircher, M., Del Campo, M., Amato, R.S., and Agarwal, A.K. (2015). 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. 10.1002/ajmg.a.37115.

77. Schrauwen, I., Szelinger, S., Siniard, A.L., Kurdoglu, A., Corneveaux, J.J., Malenica, I., Richholt, R., Van Camp, G., De Both, M., Swaminathan, S., et al. (2015). A frame-shift mutation in CAV1 is associated with a severe neonatal progeroid and lipodystrophy syndrome. PLoS One *10*, e0131797. 10.1371/journal.pone.0131797

PONE-D-14-08911 [pii].

78. Marsboom, G., Chen, Z., Yuan, Y., Zhang, Y., Tiruppathi, C., Loyd, J.E., Austin, E.D., Machado, R.F., Minshall, R.D., Rehman, J., and Malik, A.B. (2017). 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. 10.1091/mbc.E16-11-0790.

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

80. Scherer, P.E., Tang, Z., Chun, M., Sargiacomo, M., Lodish, H.F., and Lisanti, M.P. (1995). Caveolin isoforms differ in their N-terminal protein sequence and subcellular distribution. Identification and epitope mapping of an isoform-specific monoclonal antibody probe. J. Biol. Chem. *270*, 16395-16401.

81. Kogo, H., and Fujimoto, T. (2000). Caveolin-1 isoforms are encoded by distinct mRNAs. Identification of mouse caveolin-1 mRNA variants caused by alternative transcription initiation and splicing. FEBS Lett *465*, 119-123.

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

83. Kirchner, P., Bug, M., and Meyer, H. (2013). Ubiquitination of the N-terminal region of caveolin-1 regulates endosomal sorting by the VCP/p97 AAA-ATPase. J Biol Chem *288*, 7363-7372. M112.429076 [pii]

10.1074/jbc.M112.429076.

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

85. Li, S., Seitz, R., and Lisanti, M.P. (1996). 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.

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

87. Corley Mastick, C., Sanguinetti, A.R., Knesek, J.H., Mastick, G.S., and Newcomb, L.F. (2001). 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. 10.1006/excr.2001.5205.

88. Mastick, C.C., and Saltiel, A.R. (1997). Insulin-stimulated tyrosine phosphorylation of caveolin is specific for the differentiated adipocyte phenotype in 3T3-L1 cells. J Biol Chem *272*, 20706-20714. 10.1074/jbc.272.33.20706.

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

90. Doucey, M.A., Bender, F.C., Hess, D., Hofsteenge, J., and Bron, C. (2006). Caveolin-1 interacts with the chaperone complex TCP-1 and modulates its protein folding activity. Cell Mol Life Sci *63*, 939-948. 10.1007/s00018-005-5551-z.

91. Busch, D.J., Houser, J.R., Hayden, C.C., Sherman, M.B., Lafer, E.M., and Stachowiak, J.C. (2015). Intrinsically disordered proteins drive membrane curvature. Nat Commun *6*, 7875. 10.1038/ncomms8875.

92. Yuan, F., Alimohamadi, H., Bakka, B., Trementozzi, A.N., Day, K.J., Fawzi, N.L., Rangamani, P., and Stachowiak, J.C. (2021). Membrane bending by protein phase separation. Proc Natl Acad Sci U S A *118*, e2017435118. 10.1073/pnas.2017435118.

93. Tillu, V.A., Rae, J., Gao, Y., Ariotti, N., Floetenmeyer, M., Kovtun, O., McMahon, K.A., Chaudhary, N., Parton, R.G., and Collins, B.M. (2021). Cavin1 intrinsically disordered domains are essential for fuzzy electrostatic interactions and caveola formation. Nat Commun *12*, 931. 10.1038/s41467-021-21035-4.

94. King, N., Hittinger, C.T., and Carroll, S.B. (2003). Evolution of key cell signaling and adhesion protein families predates animal origins. Science *301*, 361-363. 10.1126/science.1083853.

95. Segawa, Y., Suga, H., Iwabe, N., Oneyama, C., Akagi, T., Miyata, T., and Okada, M. (2006). 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. 10.1073/pnas.0600021103.

96. Parton, R.G., Del Pozo, M.A., Vassilopoulos, S., Nabi, I.R., Le Lay, S., Lundmark, R., Kenworthy, A.K., Camus, A., Blouin, C.M., Sessa, W.C., and Lamaze, C. (2020). Caveolae: The FAQs. Traffic *21*, 181-185. 10.1111/tra.12689.

97. Walser, P.J., Ariotti, N., Howes, M., Ferguson, C., Webb, R., Schwudke, D., Leneva, N., Cho, K.J., Cooper, L., Rae, J., et al. (2012). Constitutive formation of caveolae in a bacterium. Cell *150*, 752-763. S0092-8674(12)00886-0 [pii]

10.1016/j.cell.2012.06.042.

98. Booth, D.S., Szmidt-Middleton, H., and King, N. (2018). Transfection of choanoflagellates illuminates their cell biology and the ancestry of animal septins. Mol Biol Cell *29*, 3026-3038. 10.1091/mbc.E18-08-0514.

99. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191. 10.1093/bioinformatics/btp033.

100. Eddy, S.R. (2011). Accelerated Profile HMM Searches. PLoS Comput Biol *7*, e1002195. 10.1371/journal.pcbi.1002195.

101. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res *30*, 3059-3066. 10.1093/nar/gkf436.

102. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780. 10.1093/molbev/mst010.

103. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics *27*, 1164-1165. 10.1093/bioinformatics/btr088.

104. Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688-2690. 10.1093/bioinformatics/btl446.

105. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313. 10.1093/bioinformatics/btu033.

106. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526-528. 10.1093/bioinformatics/bty633.

Figure 1

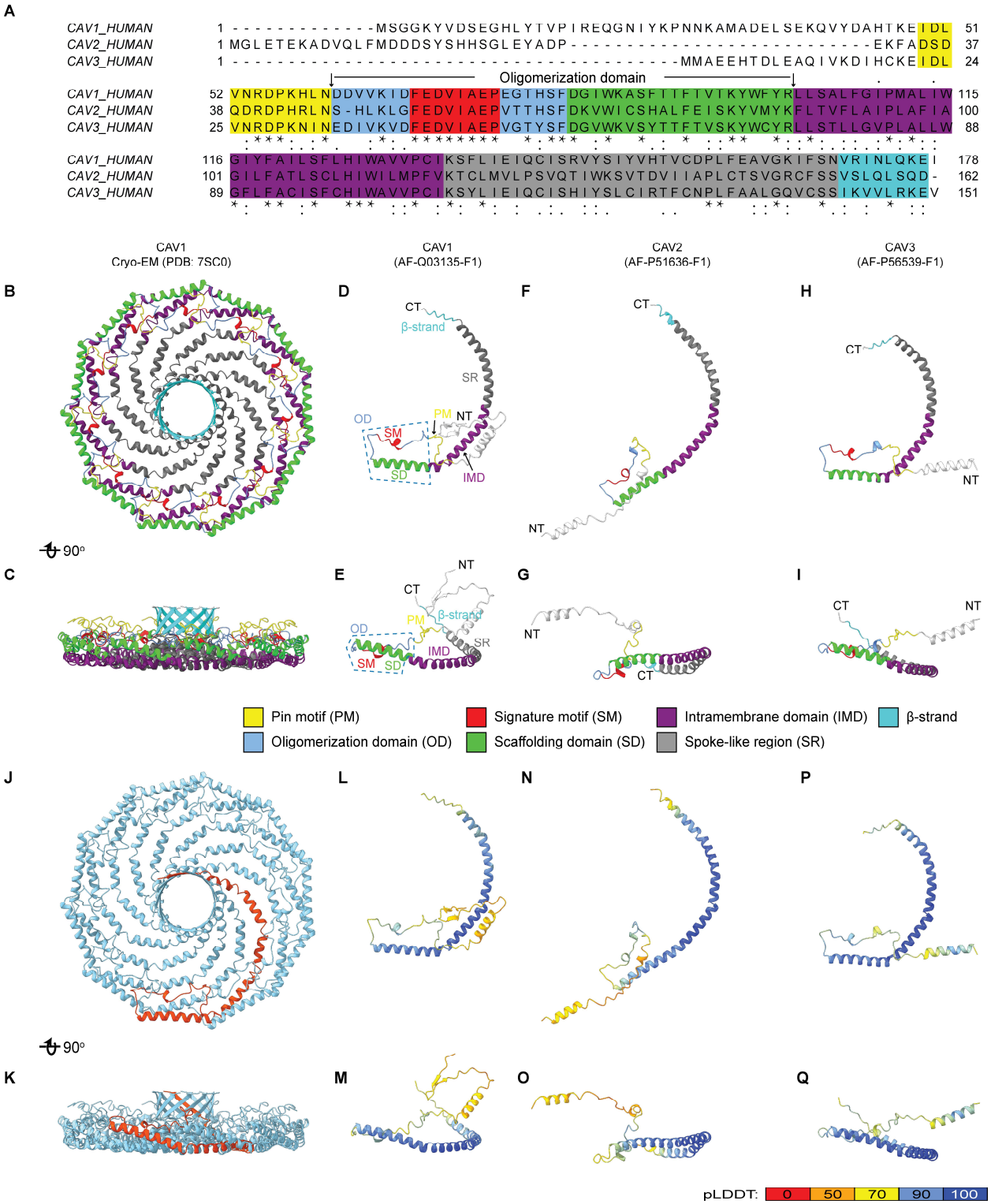

A
CAV1_HUMAN
CAV2_HUMAN
CAV3_HUMAN
Oligomerization domain
CAV1
Cryo-EM (PDB: 7SC0)
CAV1
(AF-Q03135-F1)
CAV2
(AF-P51636-F1)
CAV3
(AF-P56539-F1)
B
C
D
E
F
G
H
I
CT
NT
β-strand
SR
OD
PM
SM
SD
IMD
90°
Pin motif (PM)
Signature motif (SM)
Intramembrane domain (IMD)
β-strand
Oligomerization domain (OD)
Scaffolding domain (SD)
Spoke-like region (SR)
J
K
L
M
N
O
P
Q
pLDDT:
0
50
70
90
100

Figure 2

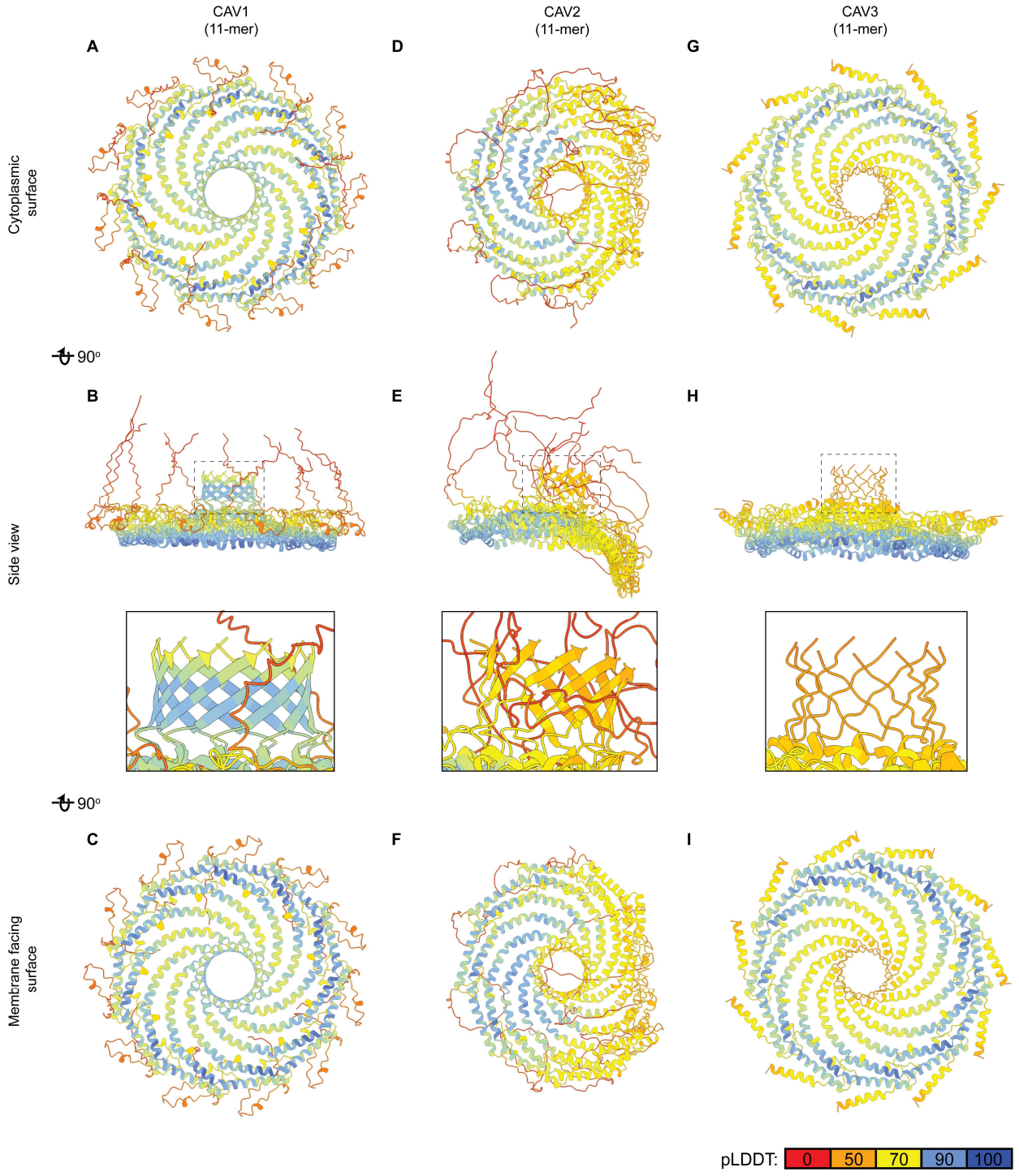

CAV1
(11-mer)
CAV2
(11-mer)
CAV3
(11-mer)
A
D
G
Cytoplasmic surface
90°
B
E
H
Side view
90°
C
F
I
Membrane facing surface
pLDDT:
0
50
70
90
100

Figure 3

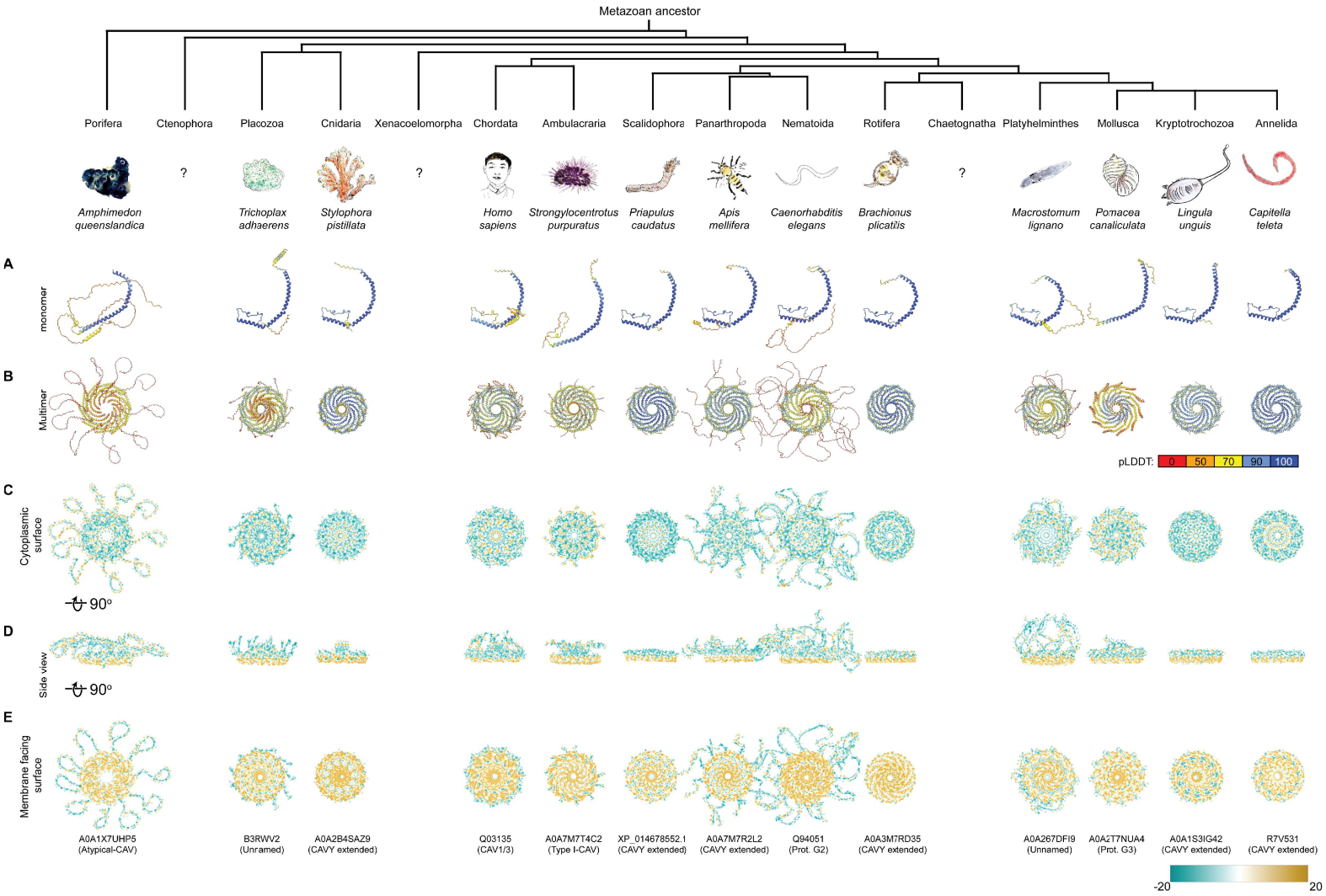

Metazoan ancestor
Porifera
Ctenophora
Placozoa
Cnidaria
Xenacoelomorpha
Chordata
Ambulacraria
Scalidophora
Panarthropoda
Nematoida
Rotifera
Chaetognatha
Platyhelminthes
Mollusca
Kryptotrochozoa
Annelida
Amphimedon queenslandica
Trichoplax adhaerens
Stylophora pistillata
Homo sapiens
Strongylocentrotus purpuratus
Priapulus caudatus
Apis mellifera
Caenorhabditis elegans
Brachionus plicatilis
Macrostomum lignano
Pomacea canaliculata
Lingula unguis
Capitella teleta
A
monomer
B
Multimer
pLDDT: 0 50 70 90 100
C
Cytoplasmic surface
90°
D
Side view
90°
E
Membrane facing surface
A0A1X7UHP5 (Atypical-CAV)
B3RWV2 (Unnamed)
A0A2B4SAZ9 (CAVY extended)
Q03135 (CAV1/3)
A0A7M7T4C2 (Type I-CAV)
XP_014678552.1 (CAVY extended)
A0A7M7R2L2 (CAVY extended)
Q94051 (Prot. G2)
A0A3M7RD35 (CAVY extended)
A0A267DFI9 (Unnamed)
A0A2T7NUA4 (Prot. G3)
A0A1S3IG42 (CAVY extended)
R7V531 (CAVY extended)
-20
20

Figure 4

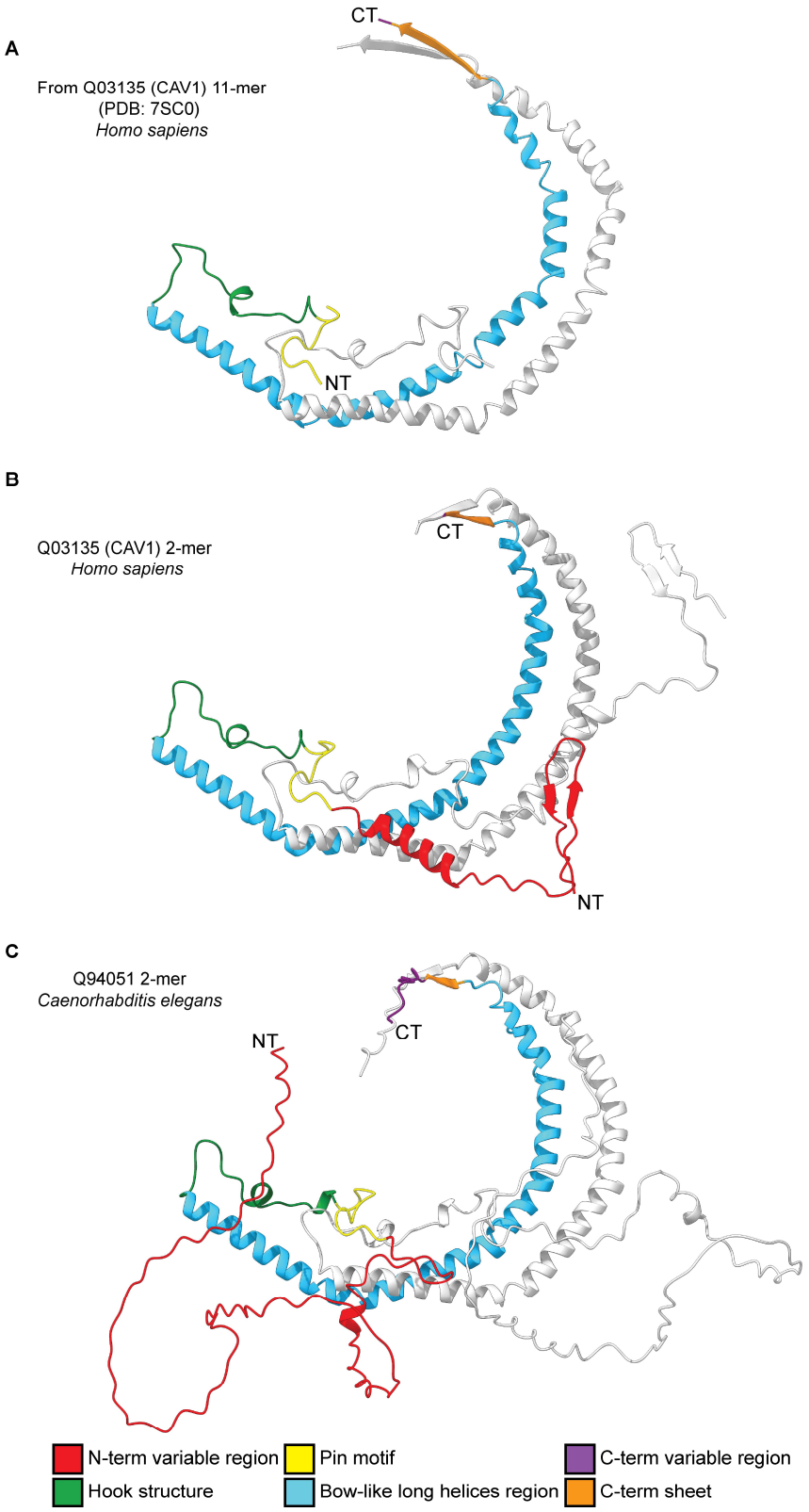

A
From Q03135 (CAV1) 11-mer
(PDB: 7SC0)
Homo sapiens
CT
NT
B
Q03135 (CAV1) 2-mer
Homo sapiens
CT
NT
C
Q94051 2-mer
Caenorhabditis elegans
NT
CT
N-term variable region
Pin motif
C-term variable region
Hook structure
Bow-like long helices region
C-term sheet

Figure 5

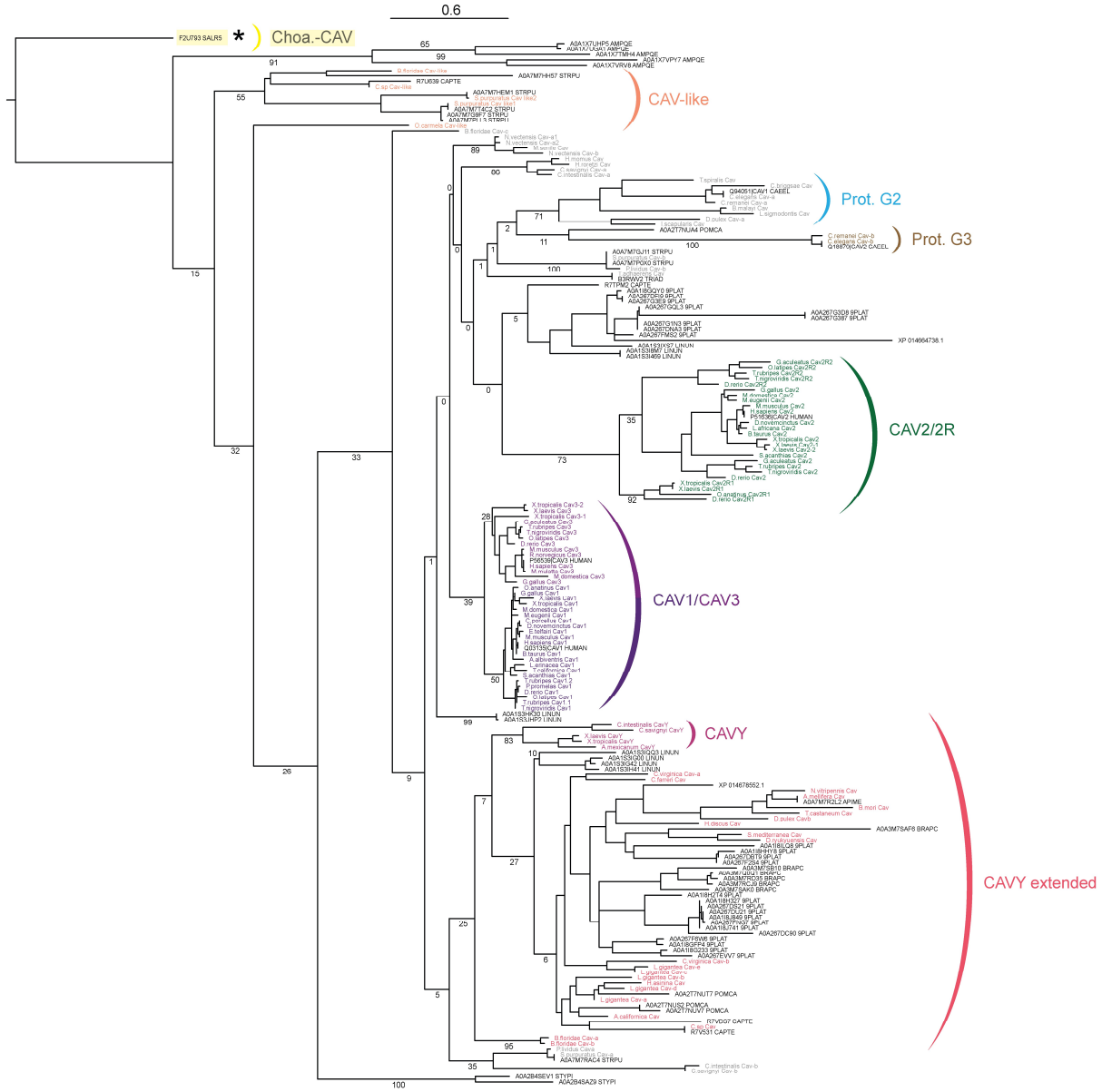

0.6
Choa.-CAV
CAV-like
Prot. G2
Prot. G3
CAV2/2R
CAV1/CAV3
CAVY
CAVY extended

Figure 6

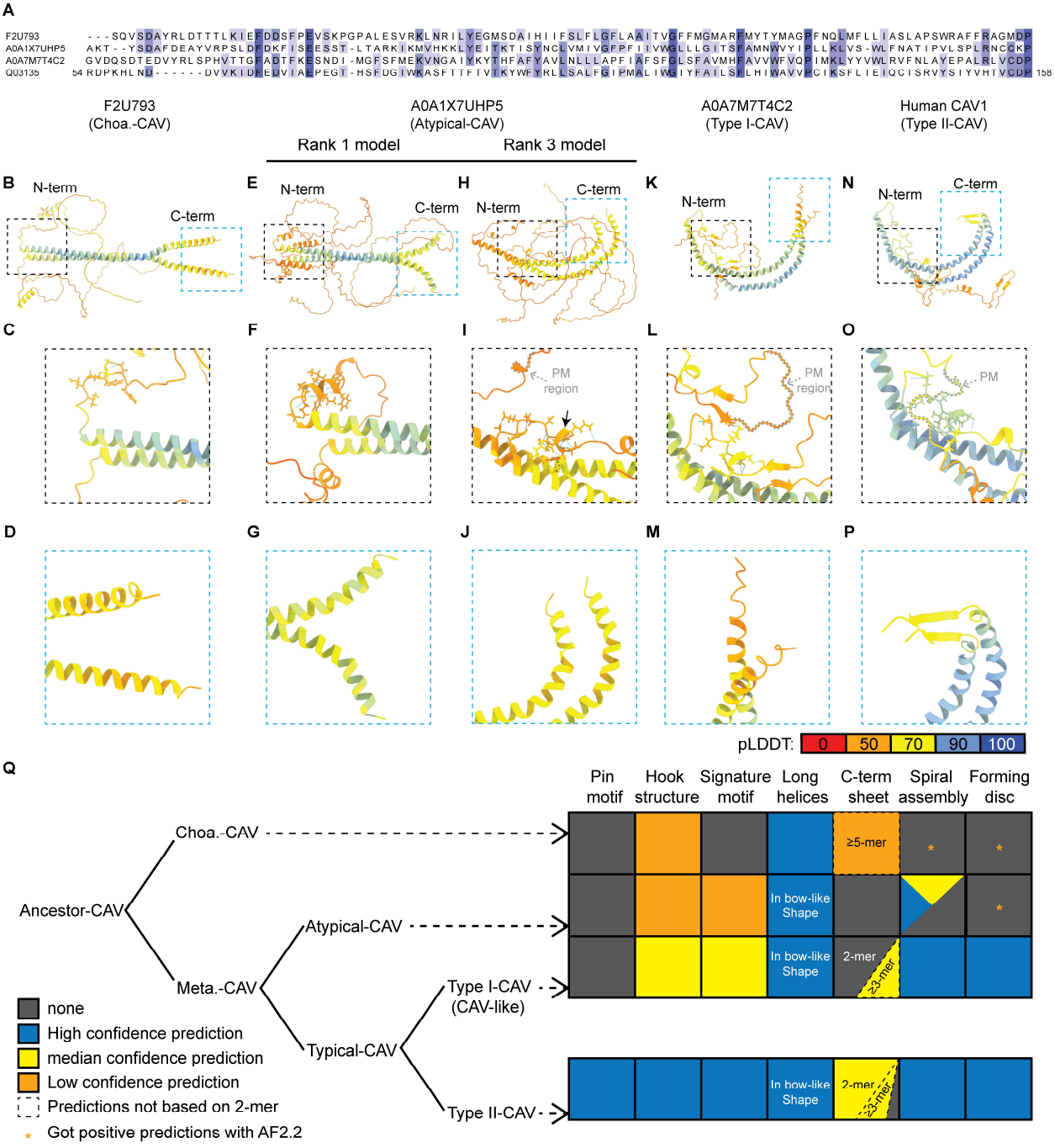

A
F2U793
A0A1X7UHP5
A0A7M7T4C2
Q03135
F2U793 (Choa.-CAV)
A0A1X7UHP5 (Atypical-CAV)
A0A7M7T4C2 (Type I-CAV)
Human CAV1 (Type II-CAV)
Rank 1 model
Rank 3 model
N-term
C-term
PM region
PM
pLDDT:
0
50
70
90
100
Q
Pin motif
Hook structure
Signature motif
Long helices
C-term sheet
Spiral assembly
Forming disc
Choa.-CAV
Ancestor-CAV
Atypical-CAV
Meta.-CAV
Type I-CAV (CAV-like)
Typical-CAV
Type II-CAV
≥5-mer
In bow-like Shape
2-mer
≥3-mer
none
High confidence prediction
median confidence prediction
Low confidence prediction
Predictions not based on 2-mer
Got positive predictions with AF2.2

Figure 7

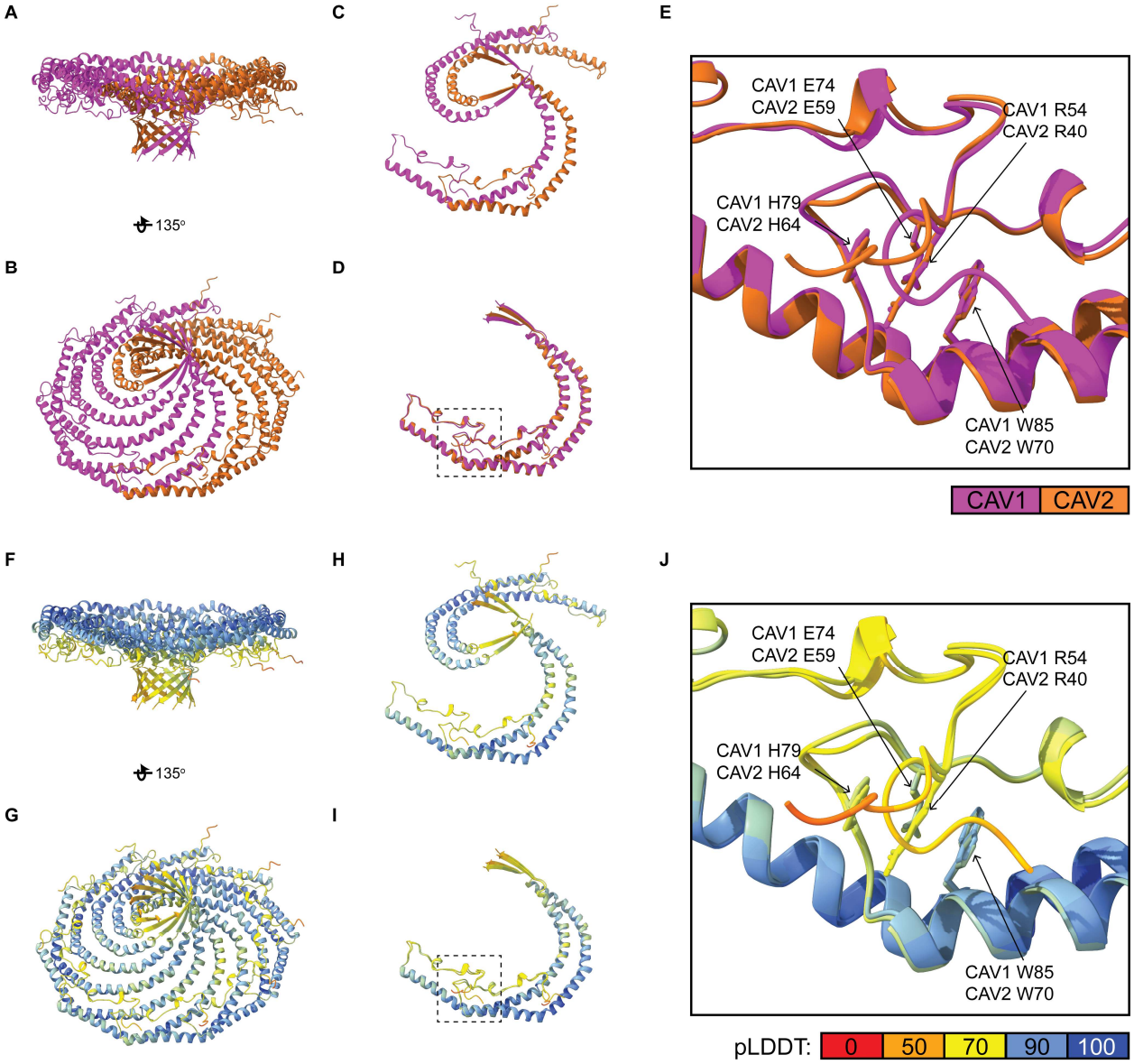
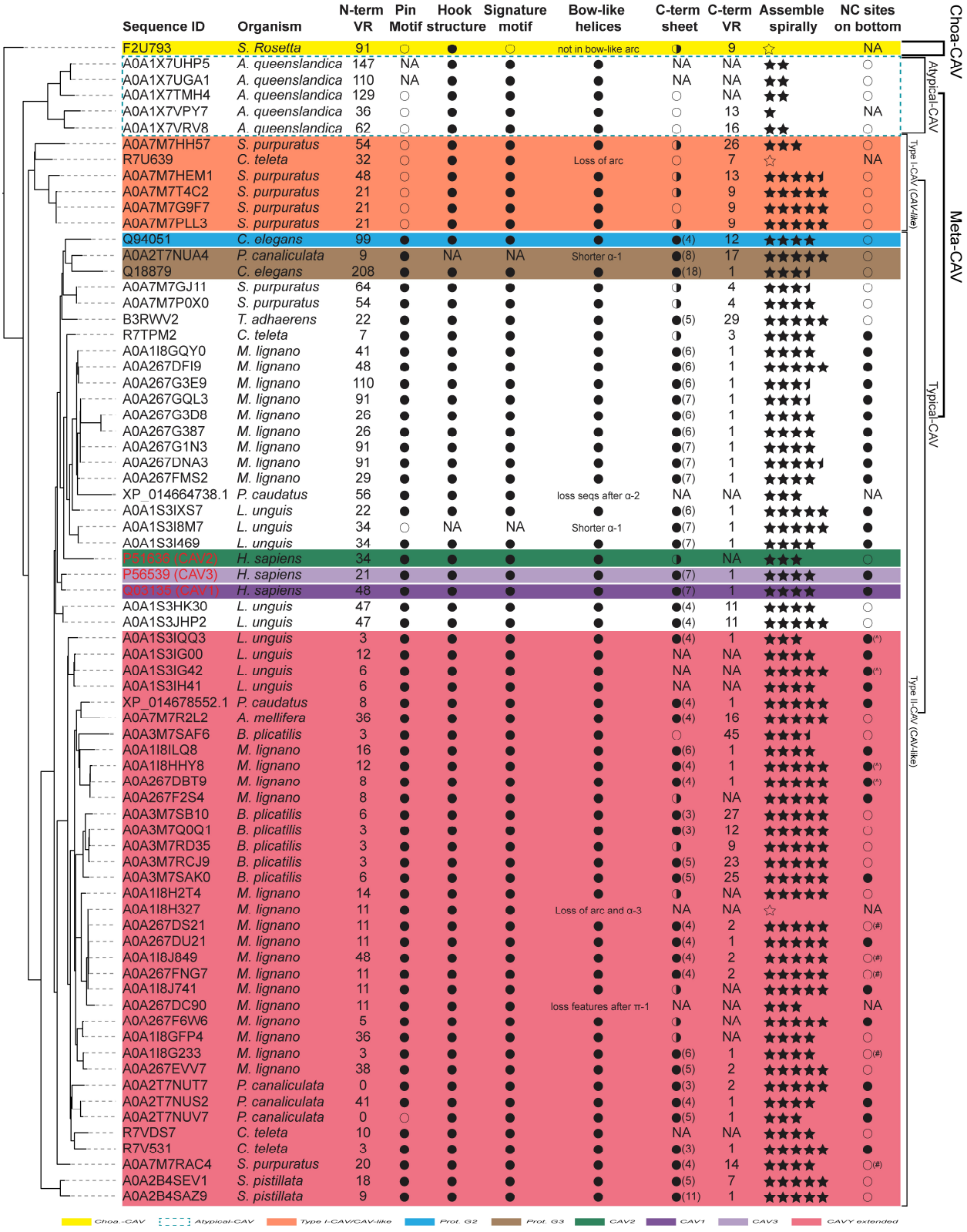

A
135°
B
C
D
E
CAV1 E74
CAV2 E59
CAV1 R54
CAV2 R40
CAV1 H79
CAV2 H64
CAV1 W85
CAV2 W70
CAV1
CAV2
F
135°
G
H
I
J
CAV1 E74
CAV2 E59
CAV1 R54
CAV2 R40
CAV1 H79
CAV2 H64
CAV1 W85
CAV2 W70
pLDDT:
0
50
70
90
100

Table 1

Sequence ID	Organism	N-term VR	Pin Motif	Hook structure	Signature motif	Bow-like helices	C-term sheet	C-term VR	Assemble spirally	NC sites on bottom
F2U793	*S. Rosetta*	91	○	●	○	not in bow-like arc	◑	9	☆	NA
A0A1X7UHP5	*A. queenslandica*	147	NA	●	●	●	NA	NA	★★	○
A0A1X7UGA1	*A. queenslandica*	110	NA	●	●	●	NA	NA	★★	○
A0A1X7TMH4	*A. queenslandica*	129	○	●	●	●	○	NA	★★	○
A0A1X7VPY7	*A. queenslandica*	36	○	●	●	●	○	13	★	NA
A0A1X7VRV8	*A. queenslandica*	62	○	●	●	●	○	16	★★	○
A0A7M7HH57	*S. purpuratus*	54	○	●	●	●	◑	26	★★★	○
R7U639	*C. teleta*	32	○	●	●	Loss of arc	○	7	☆	NA
A0A7M7HEM1	*S. purpuratus*	48	○	●	●	●	◑	13	★★★★½	○
A0A7M7T4C2	*S. purpuratus*	21	○	●	●	●	◑	9	★★★★★	○
A0A7M7G9F7	*S. purpuratus*	21	○	●	●	●	○	9	★★★★★	○
A0A7M7PLL3	*S. purpuratus*	21	○	●	●	●	◑	9	★★★★★	○
Q94051	*C. elegans*	99	●	●	●	●	●(4)	12	★★★★	○
A0A2T7NUA4	*P. canaliculata*	9	●	NA	NA	Shorter α-1	●(8)	17	★★★★★	○
Q18879	*C. elegans*	208	●	●	●	●	●(18)	1	★★★½	○
A0A7M7GJ11	*S. purpuratus*	64	●	●	●	●	◑	4	★★★½	○
A0A7M7P0X0	*S. purpuratus*	54	●	●	●	●	◑	4	★★★★	○
B3RWV2	*T. adhaerens*	22	●	●	●	●	●(5)	29	★★★★★	○
R7TPM2	*C. teleta*	7	●	●	●	●	◑	3	★★★★	●
A0A1I8GQY0	*M. lignano*	41	●	●	●	●	●(6)	1	★★★★	●
A0A267DFI9	*M. lignano*	48	●	●	●	●	●(6)	1	★★★★★	●
A0A267G3E9	*M. lignano*	110	●	●	●	●	●(6)	1	★★★½	●
A0A267GQL3	*M. lignano*	91	●	●	●	●	●(7)	1	★★★½	●
A0A267G3D8	*M. lignano*	26	●	●	●	●	●(6)	1	★★★★	●
A0A267G387	*M. lignano*	26	●	●	●	●	●(6)	1	★★★★	●
A0A267G1N3	*M. lignano*	91	●	●	●	●	●(7)	1	★★★★	●
A0A267DNA3	*M. lignano*	91	●	●	●	●	●(7)	1	★★★★½	●
A0A267FMS2	*M. lignano*	29	●	●	●	●	●(7)	1	★★★★	●
XP_014664738.1	*P. caudatus*	56	●	●	●	loss seqs after α-2	NA	NA	★★★	NA
A0A1S3IXS7	*L. unguis*	22	●	●	●	●	●(6)	1	★★★★★	●
A0A1S3I8M7	*L. unguis*	34	○	NA	NA	Shorter α-1	●(7)	1	★★★★★	●
A0A1S3I469	*L. unguis*	34	●	●	●	●	●(7)	1	★★★★	●
P56165A (CAV2)	*H. sapiens*	34	●	●	●	●	◑	NA	★★★	○
P56539 (CAV3)	*H. sapiens*	21	●	●	●	●	●(7)	1	★★★★	●
Q03135 (CAV1)	*H. sapiens*	48	●	●	●	●	●(7)	1	★★★★	●
A0A1S3HK30	*L. unguis*	47	●	●	●	●	●(4)	11	★★★★	○
A0A1S3JHP2	*L. unguis*	47	●	●	●	●	●(4)	11	★★★★★	○
A0A1S3IQQ3	*L. unguis*	3	●	●	●	●	●(4)	1	★★★	●(^)
A0A1S3IG00	*L. unguis*	12	●	●	●	●	NA	NA	★★★★	●
A0A1S3IG42	*L. unguis*	6	●	●	●	●	NA	NA	★★★★★	●(^)
A0A1S3IH41	*L. unguis*	6	●	●	●	●	NA	NA	★★★★	●
XP_014678552.1	*P. caudatus*	8	●	●	●	●	●(4)	1	★★★★★	●
A0A7M7R2L2	*A. mellifera*	36	●	●	●	●	●(4)	16	★★★★★	○
A0A3M7SAF6	*B. plicatilis*	3	●	●	●	●	○	45	★★★½	○
A0A1I8ILQ8	*M. lignano*	16	●	●	●	●	●(6)	1	★★★★	●
A0A1I8HHY8	*M. lignano*	12	●	●	●	●	●(4)	1	★★★★★	●(^)
A0A267DBT9	*M. lignano*	8	●	●	●	●	●(4)	1	★★★★★	●(^)
A0A267F2S4	*M. lignano*	8	●	●	●	●	◑	NA	★★★★★	●
A0A3M7SB10	*B. plicatilis*	6	●	●	●	●	●(3)	27	★★★★★	○
A0A3M7Q0Q1	*B. plicatilis*	3	●	●	●	●	●(3)	12	★★★★★	○
A0A3M7RD35	*B. plicatilis*	3	●	●	●	●	◑	9	★★★★★	○
A0A3M7RCJ9	*B. plicatilis*	3	●	●	●	●	●(5)	23	★★★★★	○
A0A3M7SAK0	*B. plicatilis*	6	●	●	●	●	●(5)	25	★★★★★	●
A0A1I8H2T4	*M. lignano*	14	●	●	●	●	◑	NA	★★★★★	○
A0A1I8H327	*M. lignano*	11	●	●	●	Loss of arc and α-3	NA	NA	☆	NA
A0A267DS21	*M. lignano*	11	●	●	●	●	●(4)	2	★★★★★	○(#)
A0A267DU21	*M. lignano*	11	●	●	●	●	●(4)	1	★★★★★	●
A0A1I8J849	*M. lignano*	48	●	●	●	●	●(4)	2	★★★★★	○(#)
A0A267FNG7	*M. lignano*	11	●	●	●	●	●(4)	2	★★★★★	○(#)
A0A1I8J741	*M. lignano*	11	●	●	●	●	◑	NA	★★★★★	●
A0A267DC90	*M. lignano*	11	●	●	●	loss features after π-1	NA	NA	★★★	NA
A0A267F6W6	*M. lignano*	5	●	●	●	●	◑	NA	★★★★★	●
A0A1I8GFP4	*M. lignano*	36	●	●	●	●	◑	NA	★★★★	○
A0A1I8G233	*M. lignano*	3	●	●	●	●	●(6)	1	★★★★	○(#)
A0A267EVV7	*M. lignano*	38	●	●	●	●	●(5)	2	★★★★★	○
A0A2T7NUT7	*P. canaliculata*	0	●	●	●	●	●(3)	2	★★★★★	●
A0A2T7NUS2	*P. canaliculata*	41	●	●	●	●	●(4)	1	★★★★	●
A0A2T7NUV7	*P. canaliculata*	0	○	●	●	●	●(5)	1	★★★	●
R7VDS7	*C. teleta*	10	●	●	●	●	NA	NA	★★★★	○
R7V531	*C. teleta*	3	●	●	●	●	●(3)	1	★★★★★	●
A0A7M7RAC4	*S. purpuratus*	20	●	●	●	●	●(4)	14	★★★★	○(#)
A0A2B4SEV1	*S. pistillata*	18	●	●	●	●	●(5)	7	★★★★★	○
A0A2B4SAZ9	*S. pistillata*	9	●	●	●	●	●(11)	1	★★★★★	○