

1    **9Å structure of the COPI coat reveals that the Arf1 GTPase occupies two**  
2    **contrasting molecular environments**

3

4    **Authors:** Svetlana O. Dodonova<sup>1,5,6</sup>, Patrick Aderhold<sup>2,6</sup>, Juergen Kopp<sup>2</sup>, Iva Ganeva<sup>2</sup>,  
5    Simone Röhling<sup>2</sup>, Wim. J. H. Hagen<sup>1</sup>, Irmgard Sinning<sup>2</sup>, Felix Wieland<sup>2</sup>, John A. G.  
6    Briggs<sup>1,3,4\*</sup>

7

8    **Affiliations:**

9    <sup>1</sup>Structural and Computational Biology Unit, European Molecular Biology Laboratory,  
10    Meyerhofstrasse 1, 69117 Heidelberg, Germany.

11    <sup>2</sup>Heidelberg University Biochemistry Center, Heidelberg University, Im Neuenheimer  
12    Feld 328, 69120 Heidelberg, Germany.

13    <sup>3</sup>Cell Biology and Biophysics Unit, European Molecular Biology Laboratory,  
14    Meyerhofstrasse 1, 69117 Heidelberg, Germany.

15    <sup>4</sup>MRC Laboratory of Molecular Biology, Cambridge CB2 0HQ, United Kingdom.

16    <sup>5</sup>Molecular Biology Department, Max Planck Institute for Biophysical Chemistry, 37077  
17    Göttingen, Germany

18    <sup>6</sup>shared first authors

19

20    \*Correspondence to: E-mail: [john.briggs@mrc-lmb.cam.ac.uk](mailto:john.briggs@mrc-lmb.cam.ac.uk)

21

22 **Keywords:**

23 COPI coat, coated vesicles, Arf1, ArfGAP, cryo-electron tomography, subtomogram

24 averaging

25

26 **Highlights:**

27 • Structure of the COPI coat combining cryo-electron tomography and

28 crystallography

29 • ArfGAP2 interacts with one of two copies of the small GTPase Arf1

30 • A cargo-binding-regulating helix in  $\delta$ -COP interacts with the other copy of Arf1

31 • Two copies of Arf1 in the COPI coat may regulate assembly and disassembly

32

33 **Impact statement:**

34 A molecular model of the assembled COPI coat, determined by cryo-electron

35 tomography of an in vitro reconstituted budding reaction, reveals details of interactions

36 mediating coat assembly and shows the binding site of ArfGAP2.

37

38 **ABSTRACT**

39 COPI coated vesicles mediate trafficking within the Golgi apparatus and between the

40 Golgi and the endoplasmic reticulum. Assembly of a COPI coated vesicle is initiated by

41 the small GTPase Arf1 that recruits the coatamer complex to the membrane, triggering

42 polymerization and budding. The vesicle uncoats before fusion with a target membrane.

43 Coat components are structurally conserved between COPI and clathrin/adaptor

44 proteins. Using cryo-electron tomography and subtomogram averaging, we determined

45 the structure of the COPI coat assembled on membranes in vitro at 9 Å resolution. We



46 also obtained a 2.57 Å resolution crystal structure of  $\beta\delta$ -COP. By combining these  
47 structures we built a molecular model of the coat. We additionally determined the coat  
48 structure in the presence of ArfGAP proteins that regulate coat dissociation. We found  
49 that Arf1 occupies contrasting molecular environments within the coat, leading us to  
50 hypothesize that some Arf1 molecules may regulate vesicle assembly while others  
51 regulate coat disassembly.

52

## 53 INTRODUCTION

54 Coated vesicles mediate transport between intracellular organelles. Coat protein 1  
55 (COPI) coated vesicles function in retrograde trafficking between the Golgi apparatus  
56 and the Endoplasmic Reticulum (ER), and between the Golgi cisternae in both  
57 retrograde and anterograde directions (Cosson and Letourneur, 1994; Letourneur et al.,  
58 1994; Orci et al., 1997). Assembly of a COPI coated vesicle is initiated by recruitment of  
59 the small GTPase Arf1 (ADP-ribosylation factor 1) to the membrane, where it is  
60 activated by a guanine exchange factor and transitions into a GTP-bound state in which  
61 its amphipathic N0 helix is inserted into the membrane. Arf1-GTP recruits the protein  
62 complex coatomer from the cytoplasm and anchors it at the membrane. Polymerization  
63 of coatomer at the membrane assembles the COPI coat, which recruits cargo molecules  
64 and membrane machinery to the assembly site (Harter and Wieland, 1998; Jackson et  
65 al., 2012). Growth of the coat, and the resulting membrane bud, increases until a vesicle  
66 is pinched off the donor membrane. GTP hydrolysis by Arf1, activated by ArfGAPs (Arf  
67 GTPase Activating Proteins), mediates uncoating of the vesicle which subsequently  
68 fuses with its target membrane.

69 Coatomer (the COPI complex) is a heteroheptameric protein complex, consisting of  $\alpha$ ,  $\beta$ ,  
70  $\beta'$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  and  $\zeta$  subunits. Coatomer can be subdivided into an outer-coat and an adaptor  
71 subcomplex (Lowe and Kreis, 1995). The outer-coat subcomplex consists of  $\alpha$ ,  $\beta'$  and  $\epsilon$   
72 subunits.  $\alpha$ - and  $\beta'$ -COP each contain two WD40  $\beta$ -propeller domains followed by an  $\alpha$ -  
73 solenoid, a structural motif that is shared by other coat proteins such as clathrin and  
74 sec13/31 from COPII (Devos et al., 2004). The adaptor subcomplex consists of  $\beta$ ,  $\gamma$ ,  $\delta$   
75 and  $\zeta$  subunits and is homologous to the clathrin adaptor AP complexes (Schledzewski  
76 et al., 1999). Within the adaptor subcomplex,  $\beta\delta$  and  $\gamma\zeta$  are structural homologues that  
77 are likely to have arisen by gene duplication. Each of the large adaptor subunits  $\beta$ - and  
78  $\gamma$ -COP contains an  $\alpha$ -solenoid “trunk” domain, connected by an unstructured linker  
79 region to a small appendage domain. The trunk domains each interact with an Arf1  
80 molecule to anchor the coat to the membrane. In contrast to clathrin and COPII, where  
81 outer-coat and adaptor are separate complexes, in COPI they form a single protein  
82 complex in the cytoplasm that is recruited to the membrane *en bloc* (Hara-Kuge et al.,  
83 1994).

84 Insights into the structure and function of coated vesicles have been obtained by  
85 crystallography and cryo-electron microscopy (cryoEM) (Faini et al., 2012; Fotin et al.,  
86 2004; Lee and Goldberg, 2010; Shen et al., 2015; Stagg et al., 2006). We recently  
87 determined the architecture of the COPI coat assembled on the vesicle membrane in  
88 vitro using cryo-electron tomography (cryoET) and subtomogram averaging (SA)  
89 (Dodonova et al., 2015). We were able to fit homology models and available crystal  
90 structures into our cryoET structure to build a molecular model of the assembled coat.  
91 Although the COPI adaptor and outer-coat subcomplexes share multiple structural and  
92 functional similarities with the clathrin coat, they appear to be organized very  
93 differently. The COPI adaptor and outer-coat subcomplexes both form arch-like

94 arrangements on the membrane and together form a highly interconnected coat lattice:  
95 there is no clear division of function between outer-coat and adaptor subcomplexes (for  
96 consistency we will nevertheless use this nomenclature here).

97 Multiple subunits in the assembled COPI coat interact with cargo and membrane  
98 machinery. The outer-coat COPI subunits  $\alpha$  and  $\beta'$  interact with dilysine K(X)KXX cargo  
99 motifs (Jackson et al., 2012; Ma and Goldberg, 2013), while the adaptor subunit  $\gamma$   
100 interacts with p23 transmembrane peptides (Cosson and Letourneur, 1994; Harter and  
101 Wieland, 1998). COPI coated vesicles are also involved in the transport of K/HDEL  
102 cargoes associated with specific receptors (Semenza et al., 1990) that interact with the  
103 COPI coat (Majoul et al., 2001). The transport of HDEL cargoes is dependent on the  
104 presence of an amphipathic helix in  $\delta$ -COP (Arakel et al., 2016).

105 Disassembly of the coat is facilitated by ArfGAPs that induce GTP hydrolysis in Arf1 and  
106 coat disassembly using a highly conserved catalytic zinc-finger domain (Cukierman et  
107 al., 1995). Two classes, ArfGAP1 and ArfGAP2/3, have been implicated in COPI vesicular  
108 transport (Frigerio et al., 2007; Liu et al., 2005; Poon et al., 1999). ArfGAP1 contains  
109 membrane-binding ALPS (Amphipathic Lipid Packing Sensor) motifs that allow the  
110 protein to be recruited to membranes in the absence of coatomer (Bigay et al., 2005;  
111 Weimer et al., 2008). ArfGAP2 and 3 interact with COPI via their non-catalytic, C-  
112 terminal regions (Frigerio et al., 2007; Kliouchnikov et al., 2009; Yahara et al., 2006) and  
113 require coatomer for efficient membrane recruitment and activity (Szafer et al., 2001).  
114 ArfGAP2 but not ArfGAP1 was found to be associated with COPI coated vesicles budded  
115 from Golgi membranes *in vitro* (Frigerio et al., 2007). Although the functions of all three  
116 ArfGAP proteins overlap (Saitoh et al., 2009), ArfGAP2 deletion leads to a much stronger  
117 retrograde trafficking phenotype than ArfGAP1 deletion (Poon et al., 1999). Why COPI is

regulated by multiple and different ArfGAP proteins remains unclear.

## RESULTS AND DISCUSSION

### X-ray structure of $\beta$ 20-390 $\delta$ -150-COP

To understand the structure of the COPI coat, and how its assembly and disassembly are regulated, we wished to generate a complete structural model for the assembled coat at sufficient resolution to precisely position the secondary structure elements of the component proteins. Crystal structures were available for the majority of the components of the COPI coat, but not for the  $\beta\delta$ -COP subcomplex. As a first step towards a complete structural model, we expressed and crystallized a complex containing  $\beta$ -COP19-391 and  $\delta$ -COP1-175 from *C. thermophilum* (Ct), from which we obtained a molecular model of  $\beta\delta$ -COP comprising residues 20-390 of  $\beta$ -COP (“trunk region”) and residues 2-150 of  $\delta$ -COP (Figure 1, Table 1). The overall architecture of  $\beta\delta$ -COP is similar to that of the homologous  $\gamma\zeta$ -COP (Yu et al., 2012).  $\beta$ -COP(20-390) is a curved HEAT-repeat type  $\alpha$ -solenoid consisting of 20 consecutive  $\alpha$ -helices. The right-handed super-helical arch of  $\beta$ -COP wraps around the  $\delta$ -subunit. The structural model of  $\delta$ -COP comprises a longin domain followed by a  $\beta$ -turn (117-122), and two helices a (127-135) and b (139-150), which are part of the linker region connecting the longin domain and the  $\mu$ -Homology Domain (MHD). The stretch of residues between helix a and amphipathic helix b traverses the  $\beta$ -COP  $\alpha$ -solenoid in the vicinity of helix 7. Helix b, including isoleucine residues 143, 146 and 147, projects away from the  $\alpha$ -solenoid (Figure 1 – Figure Supplement 1) where it would be accessible for binding partners. These isoleucine residues have been shown to be essential for correct trafficking of HDEL motif containing cargo by COPI (Arakel et al., 2016).

142

### 143 **CryoET structure of the assembled coat**

144 We prepared COPI coated vesicles in vitro as described previously (Dodonova et al.,  
145 2015) and vitrified them for cryoET. Tomographic data was collected making use of a  
146 direct electron detector and optimized data collection conditions (Hagen et al., 2016).  
147 Subtomogram averaging was performed essentially as described in the materials and  
148 methods, and the structure was further refined by combining multiple locally-aligned  
149 structures. From 1733 vesicles and near-complete buds we obtained a structure of the  
150 leaf (representing one coatomer complex with two Arf1 molecules), the asymmetric  
151 unit of the COPI coat, at 9.2 Å resolution (Figure 2, Figure 2 – Figure Supplement 1,  
152 Video 1). At this resolution, distinct  $\alpha$ -helical densities and  $\beta$ -propeller blades are  
153 resolved (See also Figure 3A and B).

154 In order to generate a complete pseudo-atomic model of the COPI coat, we performed  
155 rigid body fitting of the available crystal structures of COPI components, including our  
156 new x-ray structure of  $\beta\delta$ -COP, into the structure of the leaf. Next we fitted the  
157 homology models of domains for which high-resolution structures are not available ( $\alpha$ -  
158 COP 327-813,  $\beta$ -COP 410-968,  $\gamma$ -COP 312-549), into the leaf EM map. It is important to  
159 note, that the COPI complex is highly conserved among eukaryotes. The sequence  
160 identity between the  $\beta\delta$ -COP from *Chaetomium thermophilum* used for crystallography  
161 and *Mus musculus* used for subtomogram averaging is 53 %. The excellent match  
162 between the fitted structures and the density further confirmed the structural  
163 conservation of the complex as well as the quality and validity of the map (see also  
164 Figure 1 Figure Supplement 1E). Flexible fitting (Trabuco et al., 2008) was performed to  
165 relax the structures and models into the map and generate an initial structural model

for the assembled coat. While the core subunits are in positions very similar to those in our previous model (Dodonova et al., 2015), there are larger changes in the positions of subunits for which only homology models were available. The improved resolution allowed us to resolve ambiguities in the orientations of appendage domains (Figure 3). Most importantly, since individual secondary structure elements are resolved, we were able to identify and assign structural elements that function separately from the folded domains (see below). The higher resolution of the EM map when compared to our previously published study (Figure 2A-B, Figure 2 – Figure Supplement 1 C-E), and the availability of the  $\beta\delta$ -COP structure, therefore allowed us to generate a more complete and significantly improved pseudo-atomic model of the coat (Figure 2C-D).

### **$\beta$ -COP and $\gamma$ -COP appendage domains**

The  $\gamma$ - and  $\beta$ -COP appendage domains are homologous to the  $\alpha$  and  $\beta 2$  ear/appendage domains of the clathrin adaptor AP-2. Both appendage domains are important for viability in yeast (DeRegis et al., 2008; Hoffman et al., 2003). In our previous model we found that the  $\gamma$ -COP appendage domain is bound to the second  $\beta$ -propeller of  $\beta'$ -COPI, while the  $\beta$ -COP appendage domain is tucked into the center of the leaf between  $\alpha$ -COP and  $\gamma$ -COP. Our new structure shows that  $\gamma$ -COP interacts with  $\beta'$ -COP via a highly conserved interface in the base of the  $\gamma$ -COP  $\beta$ -sandwich subdomain (Figure 3A, B, C). The platform subdomain of the  $\gamma$ -COP appendage is exposed (Video 1) and easily accessible for other binding partners such as ArfGAP (Watson et al., 2004). The  $\beta 2$  and  $\alpha 2$  appendage domains of AP2 are also thought to function as recruitment hubs for interaction partners, suggesting that these domains are also in an accessible location within the clathrin vesicle (Praefcke et al., 2004; Schmid et al., 2006). To date, there are

no structures available showing the positions of the AP2 appendage domains relative to AP2 or the clathrin coat.

We were previously unable to determine the orientation of the  $\beta$ -COP appendage domain. In the new structure, the  $\beta$ -sandwich and the platform subdomains of the  $\beta$ -COP appendage are clearly visible in the EM map (Video 1). The platform interacts with the trunk domain of  $\gamma$ -COP, whereas the  $\beta$ -sandwich forms a conserved interface with  $\alpha$ -COP including both  $\alpha$ -COP  $\beta$ -propeller domains and part of the  $\alpha$ -COP  $\alpha$ -solenoid domain (Figure 3D). In the clathrin system, the  $\beta$ 2-appendage domain of AP2 plays a conceptually similar role, interacting with clathrin and promoting cage formation (Owen et al., 2000). The  $\beta$ 2-appendage is thought to be linked to the core of AP2 only by its flexible linker. In contrast, the  $\beta$ -COP appendage appears to form the main link between the adaptor and the outer-coat subcomplexes of the COPI coat. The other connection is a rather small interface between the  $\beta$ -COP trunk and  $\beta'$ -COP. We speculate that the position of the  $\beta$ -COP appendage domain would allow it to modulate the conformation of coatomer by fixing the angle between the two  $\beta$ -propeller domains of  $\alpha$ -COP and by forming a rigid buttress between the adaptor subunit  $\gamma$ -COP and the outer-coat subunit  $\alpha$ -COP (Figure 3D, Video 1), thereby determining their relative arrangements. Thus, the  $\beta$ -COP appendage appears to act as a keystone in the assembled COPI coat.

## **Assignment of extra densities**

After fitting known structures and homology models to generate the initial structural model, we identified seven positions in the map where there were substantial regions of electron density not occupied by any of the fitted coatomer domains (see Methods)

(Figure 4A). We wished to identify the protein components that contribute these extra densities. We first performed secondary structure predictions for all COPI subunits using the Quick2D server (Jones, 1999; Ouali and King, 2000). We identified putative secondary structure elements in multiple COPI subunits that were outside the regions included in the crystal structures and homology models of the COPI domains. We compared the sequence positions of these elements with the locations of the unoccupied densities, as well as with available cross-linking mass-spectrometry (XL/MS) data, thereby assigning secondary structure elements to unoccupied densities.

The C-terminal region of  $\beta'$ -COP was predicted to contain two additional  $\alpha$ -helices, consistent with the presence in the EM density of elongated, unoccupied densities adjacent to the C-terminus of  $\beta'$ -COP at the  $\beta'$ - $\alpha$  interface (Figure 4A, B “1”). We speculate that these helices may stabilize or regulate flexibility at the  $\beta'$ - $\alpha$  solenoid-solenoid interface.

We interpret the unoccupied density adjacent to the N-terminal  $\beta$ -propeller of  $\alpha$ -COP, in the vicinity of the Arf1 bound to  $\beta$ -COP ( $\beta$ Arf1), as being partly contributed by a flexible hydrophobic loop of the  $\beta$ -propeller itself that was absent from the crystal structure of the propeller domain (Figure 4C “2”) (PDB:4J87, (Ma and Goldberg, 2013)). Flexibility of the loop likely results in smearing of the density, and we cannot rule out that this density also contains other protein components (e.g. unstructured linker regions).

Five additional  $\alpha$ -helices were predicted C-terminal to the  $\beta$ -trunk domain, consistent with the presence of several unoccupied densities near the C-terminal end of the fitted  $\beta$ -trunk model (Figure 4A, D “3” and “4”). Cross-links between these extra helices and the  $\zeta$ -longin domain support their location in that region ( $\beta$ 617-  $\zeta$ 39,  $\beta$ 618-  $\zeta$ 39) (Table 2). After these additional helices, the 103 amino acid unstructured  $\beta$ -COP linker extends



in the sequence until the start of the  $\beta$ -appendage domain. The theoretical fully extended length of the linker is around 370 Å. Within the assembled coat it must span a distance of approximately 100 Å. Several cross-links between the linker, the  $\beta$ -COP appendage domain,  $\beta'$ -COP, and  $\alpha$ -COP (Table 2, rows 5-15) are consistent with the expected route of the linker through the assembled coat, schematically shown in Figure 3F. The small density “6” (Figure 4A), may represent a short helical stretch in the linker bound to the  $\beta$ -trunk. Density “5” is discussed below.

#### **$\delta$ -COP helices “a, b and c”**

Secondary structure predictions indicate the presence of three putative  $\alpha$ -helices (referred to as “a, b and c”) in the linker-region between the  $\delta$ -COP longin domain and MHD (Figure 5A).

Helix a and the first part of helix b are present in our x-ray structure (Figure 1). One short and one long extended rod-like density corresponding to these helices are also observed in equivalent positions immediately C-terminal to the fitted  $\delta$ -COP longin domain in our COPI coat EM density (Figure 5B). While no electron density was visible for residues 151-175 of  $\delta$ -COP in the x-ray structure, our cryoEM map of the COPI coat has clear density corresponding to a longer helix, and we modeled residues 139-165 of the predicted helix b into our density (Figure 5).

Helix a interacts with the  $\delta$ -COP longin and the  $\beta$ -COP trunk domains. The N-terminal part of helix b interacts with the  $\beta$ -COP trunk domain, while its C-terminal part interacts with the peripheral Arf1 molecule ( $\beta$ Arf1) in the triad (Figure 5B). The very C-terminal end of helix b also contacts the membrane surface (Figure 5C). An equivalent helix b is

predicted in the sequence of the homologous  $\zeta$ -COP (Alisaraie and Rouiller, 2012), but we saw no equivalent density near  $\zeta$ -COP in our EM structure.

In an electron microscopy structure of a soluble trimeric assembly of AP-1, Nef, and Arf1 (Shen et al., 2015) an interaction was observed between the  $\mu$ 1 subunit of AP-1 (homologous to  $\delta$ -COP), and Arf1. The interaction made by  $\mu$ 1 is very different to that made by  $\delta$ -COP, since it involves the MHD of  $\mu$ 1 instead of the longin domain, and a different surface of Arf1.

Several cross-links ( $\delta$ 233- $\delta$ 263,  $\delta$ 243- $\delta$ 263,  $\delta$ 233- $\beta'$ 627 and  $\delta$ 241- $\beta'$ 627) (Table 2) suggest the approximate location of helix c (Figure 6 – Figure Supplement 1), and based on these observations we speculate that density “5” is occupied by this  $\alpha$ -helix (Figure 4A, D “5”). In this position, helix c could help to coordinate the positioning of the C-terminal  $\delta$ -COP MHD, which could otherwise be located a long distance from the vesicle: there are  $\sim$ 100 amino acids between helix b and the MHD.

### **Interactions between $\delta$ -COP and Arf1**

Further validation of the interaction between  $\delta$ -COP helix b and Arf1 is provided by published photo-cross-linking data (Sun et al., 2007), which showed a cross-link between Arf1 residue 167 and  $\delta$ -COP (Figure 6C, Arf167 is marked), as well as by mass-spectrometry cross-linking data that showed cross-links between the second longer helix b and Arf1 ( $\delta$ 142-Arf36,  $\delta$ 164-Arf36) (Table 2, Figure 6C). We prepared  $\delta$ -COP variants containing photolabile amino acids in position 156 or 159. These variants were expressed as subcomplexes with  $\beta$ -COP(19-391) and recruited to liposomes in an Arf- and GTP-dependent manner. Subsequent photo-cross-linking resulted in an 80kDa product that was confirmed by western blotting to consist of Arf1 and  $\delta$ -COP (Figure

6A). We also prepared an Arf1 variant with two photolabile amino acid derivatives in positions 46 (known to cross-link to  $\beta$ -COP (Sun et al 2007)) and 167. This Arf1 could be photo-cross-linked within vesicles to give a 180 kD product, corresponding to  $\beta$ -COP,  $\delta$ -COP and Arf1 (Figure 6B, and Figure 6 – Figure Supplement 2), confirming that one Arf1 molecule interacts with both  $\beta$ - and  $\delta$ -COP within the same coatamer complex.

To dissect the role of the  $\delta$ -COP linker-region ( $\delta$ COP 117-271) in binding to Arf1, we made four  $\delta$ -COP constructs incorporating different parts of this linker region. These terminate after helix a ( $\delta$ 1-137), helix b ( $\delta$ 1-175) or helix c ( $\delta$ 1-243) or code for full-length  $\delta$ -COP. These proteins were expressed as subcomplexes with  $\beta$ -COP(19-391) and tested for binding to Arf1 in its GMPPNP-loaded state using pulldown assays (Figure 6D). The complex containing  $\delta$ 1-137 showed very little GTP dependent Arf1 binding,  $\delta$ 1-175 showed an intermediate level, while  $\delta$ 1-243 bound Arf1 in a GTP dependent manner at the same level as the complex containing full length  $\delta$ -COP (Figure 6E). These results indicate that the interaction between  $\beta\delta$ -COP and Arf1-GTP is stabilized by  $\delta$ -COP helix b and is further stabilized by downstream regions of  $\delta$ -COP including helix c. These observations are consistent with our structural model for the assembled coat in which helix b interacts directly with Arf1, and helix c is bound to  $\beta$ -COP to stabilize the complex.

#### **A role of $\delta$ -COP in regulating coat assembly**

$\delta$ -COP helix a is analogous to the  $\alpha$ 5 helix in the AP2  $\mu$ 2-subunit linker, and they are positioned similarly near the  $\beta$ -COP/ $\beta$ 2-AP trunk. Helix b is a conserved, COPI-specific feature, at a sequence distance similar to that of the “bind-back” helix in AP2  $\mu$ 2 (Arakel et al., 2016; Jackson et al., 2010). Arakel and colleagues proposed two models to explain

a requirement of helix b for retrieval of HDEL/KDEL proteins (Arakel et al., 2016). In a first model, this helix was proposed to bind back into a furrow in  $\beta$ -COP, thereby preventing destabilization of the  $\beta$ -COP  $\alpha$ -solenoid. In a second model, helix b, which is amphipathic in nature, binds to the membrane, bringing coatomer into close proximity where it can more easily interact with retrieval signals. Our EM structure confirms that helix b is located proximally to the membrane (Figure 5C), however the contact with the membrane is small and involves only the C-terminal end of the helix. The distance between the membrane and the nearest conserved hydrophobic residue is too large for a direct contact. This argues against it being a conventional membrane-inserting amphipathic helix (see figure 5C). Our structure suggests that the main interaction partner of helix b is Arf1.

$\delta$ -COP helix b interacts directly with the Arf1 Switch I region, Interswitch and C-terminal helix (Figure 5B) (in contrast to the interactions observed between AP1  $\mu$ 1 subunit and Arf1, which do not involve the switch regions of Arf1 (Shen et al., 2015)).  $\delta$ -COP helix b binds the surface of Arf1 in a region that is accessible when Arf1 is in the GTP state and the amphipathic N0 helix is inserted into the membrane, but is occupied by the N0 helix when Arf1 is in the GDP state (PDB: 1MR3, (Amor et al., 1994)) (Figure 5D). The exposure of this region in Arf1-GTP when N0 moves to insert into the membrane may contribute to the nucleotide-dependent recruitment of coatomer. These interactions also reconcile the apparent nucleotide independence of the interaction between isolated  $\delta$ -COP and N $\Delta$ 17Arf1 in solution (Sun et al., 2007) with the nucleotide dependence of the interaction between  $\delta$ -COP and Arf1 during coatomer recruitment to membranes (Figure 6B): the truncation of the N0 helix in N $\Delta$ 17Arf1 exposes the  $\delta$ -COP binding site regardless of Arf's nucleotide state.

These observations suggest a mechanistic model for the role of  $\delta$ -COP helix b in COPI function. Upon binding of the Arf1 N0 amphipathic helix to the membrane, the binding site for  $\delta$ -COP is exposed, and the resulting interaction contributes directly to Arf1-dependent coat recruitment to the membrane. The interaction of  $\delta$ -COP may additionally stabilize Arf1 in its active GTP-bound state by binding its Switch and Interswitch regions. The interaction of  $\delta$ -COP helix b with the Arf1 GTPase, with the membrane, and possibly with cargo confirms and explains its critical role in regulating COPI function.

We note a further implication of this model: GTP-hydrolysis by Arf1 can directly modulate the conformation of  $\delta$ -COP. Such modulation may be transmitted to other binding partners such as the KDEL cargo receptor, providing a possible route to link coat hydrolysis to cargo binding/release.

### **The linkages between triads**

We also determined the structures of the linkages between triads as previously described (Dodonova et al., 2015) (Figure 2 – Figure Supplement 2). The linkage structures are at lower resolution than that of the leaf, so we fitted our leaf structure as a rigid body (see Materials and Methods and Figure 2 – Figure Supplement 2) to assess which subunits are close to one another in the linkages. Consistent with our previously published structures (Dodonova et al., 2015) the central contacts in linkages I and IV are made by the  $\alpha$  $\epsilon$ -COP subunits, and those in linkage II are made by the  $\delta$ -COP MHD. Interestingly neither  $\epsilon$ -COP nor the  $\delta$ -COP MHD are essential for COPI function and the combined deletion of both is not lethal for yeast (Arakel et al., 2016; Kimata et al.,

2000). While the contacts formed by these proteins at the linkages may play a regulatory role, they are therefore not essential for COPI function.

The slightly improved resolution of the linkage structures obtained here compared to our previously published maps showed that the density that we interpret as being contributed by a flexible hydrophobic loop of the N-terminal  $\beta$ -propeller of  $\alpha$ -COP, is in the vicinity of  $\gamma$ -COP in a neighbouring triad. We found that  $\beta$ Arf1 is close to the N-terminal  $\beta$ -propeller of  $\beta'$ -COP in a neighbouring triad, and approaches the trunk domain of  $\beta$ -COP (Figure 2 – Figure Supplement 2 C). None of these contacts are similar to the previously described crystal contact between the  $\alpha$ 1 subunit of AP1 and the back side of  $\beta$ Arf1 (Ren et al., 2013) suggesting that an equivalent interaction is not relevant within the COPI coat. While the low-resolution of the linkage structures precludes more detailed interpretation, the proximity of the peripheral  $\beta$ Arf1 molecule to neighbouring triads suggests that  $\beta$ Arf1 could modulate inter-triad interactions.

### **Interaction of the coat with ArfGAPs**

The  $\gamma$ Arf1 and  $\beta$ Arf1 molecules have different interaction partners within the coat ( $\gamma$ -COP, and  $\delta$ -COP and  $\beta$ -COP, respectively) and are in very different molecular environments (Figure 2D). These observations suggest that the two different Arf1 molecules in the coat are differentially regulated. To further investigate the regulation of Arf1 in the coat, we incubated COPI coated vesicles with ArfGAP1 or ArfGAP2 in the presence of a non-hydrolysable GTP analogue and determined their structures. Previously published biochemical data indicates that almost no ArfGAP1 binds to COPI vesicles generated in vitro in the presence of poorly hydrolysable GTP analogs, while ArfGAP2 is abundant in such vesicle fractions (Frigerio et al., 2007). Note, that in the

presence of hydrolysable nucleotide GTP, ArfGAP1, 2 and 3 facilitate efficient vesicle uncoating (Weimer et al., 2008). We were unable to identify any ArfGAP1 bound to vesicles produced in the presence of GTP $\gamma$ S, while in vesicles incubated with ArfGAP2 we identified an additional density near the central  $\gamma$ Arf1 molecules in a subset of COPI leaves (Figure 7A, B and Figure 7 – Figure Supplement 1). No significant additional density was observed near the peripheral  $\beta$ Arf1 molecules or at any other position in the structure. The size and shape of the additional mass corresponded very well to the catalytic domain of the ArfGAP2 protein, which we fitted into the density (Figure 7C). To further validate this fit, we superimposed Arf6 from a structure in which it had been co-crystallized with the catalytic domain of the ArfGAP homologue ASAP3 (PDB:3LVQ, (Ismail et al., 2010)), with Arf1 in our coat structure. The resulting position of the catalytic domain of ASAP3 coincided with the additional density that we observe (compare figure 7 and Figure 7 – Figure Supplement 1 C). In this structure, the catalytic domain of ArfGAP2 is positioned directly near the Arf1 nucleotide-site, where it can provide the Arginine “finger” essential for stimulation of GTP-hydrolysis. The observed position of the ArfGAP catalytic domain differs from a previous ArfGAP1-Arf1 structure (Goldberg, 1999) in which the ArfGAP1 catalytic domain is distant from the nucleotide-site of Arf1 (Figure 7 – Figure Supplement 1 D).

We were not able to resolve the C-terminal, non-catalytic part of ArfGAP2, known to be involved in coatomer binding via interactions with the  $\gamma$ -COP appendage domain (Kliouchnikov et al., 2009; Watson et al., 2004), most likely because the non-catalytic part of the ArfGAP2 protein is largely disordered and may not form a globular domain that can be positioned at 12 Å resolution (Pietrosemoli et al., 2013).

The ArfGAP2 catalytic domain is bound into a niche in the assembled coat formed by Arf1, the  $\gamma$ - and  $\beta$ -COP adaptor subunits, and the  $\beta'$ -COP outer-coat subunit (Figure 7C and D). ArfGAP2 activity is increased by 100-fold in the presence of coatomer (Luo et al., 2009; Szafer et al., 2001). This stimulation requires the presence of both the adaptor subcomplex and outer-coat subcomplex (Pevzner et al., 2012), indicating that multiple ArfGAP-coatomer interactions are functionally important. Within a triad, an ArfGAP2-binding niche is formed by  $\beta'$ -COP and  $\beta$ -COP from one leaf and by  $\gamma$ -COP from a neighboring leaf (Figure 7D), with the central  $\gamma$ Arf1 molecules at the bottom. Thus, the complete binding site for the ArfGAP2 catalytic domain is formed when the coat is assembled. This suggests a possible proofreading mechanism: ArfGAP2 recruitment, and the resulting GTP-hydrolysis and coat dissociation can only occur once the coat is assembled, minimizing premature dissociation of coatomer from the membrane.

$\beta$ Arf1 molecules in the triad do not provide an equivalent binding niche for the catalytic domain. We did not observe ArfGAP2 bound near the  $\beta$ Arf1 molecules. This is consistent with yeast-two hybrid experiments which showed an interaction between the ArfGAP2 Glo3 and  $\beta'$ -COP and  $\gamma$ -COP but not with other coatomer subunits (Eugster et al., 2000). The membrane surface is more exposed near the  $\beta$ Arf1 molecules; we suggest that this may facilitate interaction with ArfGAP1 recruited directly to the membrane via its ALPS domains.

## Summary

By combining the  $\beta\delta$ -COP crystal structure and the in vitro EM structure of the COPI coat on the vesicle membrane, we have generated a model that reveals molecular details of the coat at the level of protein secondary structure, allowing precise



positioning of protein domains and the interpretation of isolated secondary structure elements. The resulting structural model has provided novel functional insight and offers a basis for future concept-driven investigation of molecular mechanisms that underlie vesicular transport.

The two halves of the COPI adaptor subcomplex,  $\gamma\zeta$ -COP and  $\beta\delta$ -COP, are thought to have evolved by gene duplication. The structure reveals that they have functionally diverged in two important ways. Firstly, the appendage domains have divergent functions - the  $\gamma$ -appendage sits at the outside of the coat where it provides a binding site for regulatory factors. The  $\beta$ -COP appendage domain links the adaptor and the outer-coat COPI subunits within one coatomer molecule. This functional divergence seems to be mirrored in AP-1 and AP-2, although the details of the interactions are different. Secondly, the two halves of the adaptor subcomplex recruit and position Arf1 in two different regulatory environments – at the center of the triad, bound to  $\gamma$ -COP ( $\gamma$ Arf1), and at the periphery of the triad bound to  $\beta$ -COP and  $\delta$ -COP ( $\beta$ Arf1). Our structure shows that the two Arf1 molecules can be differentially regulated both during coatomer recruitment and coat disassembly -  $\beta$ Arf1 by interactions of its Switch 1 and N0 helix regions with the essential helix b downstream of the longin domain in  $\delta$ -COP, and  $\gamma$ Arf1 by recruitment of ArfGAP2. Differential regulation would allow fine-tuning of coat assembly and cargo recruitment. We speculate that  $\beta$ Arf1 molecules may be primarily regulators of coat assembly – modulating coatomer recruitment, dissociation, and the interactions between triads that may be influenced by the presence of cargo, and in which the  $\delta$ -COP helix b plays a key role. ArfGAP1, bound to the membrane, may function at this stage. We speculate that the  $\gamma$ Arf1 molecules, via their interaction with ArfGAP2, are the primary regulators of coat disassembly, unlocking the triad and triggering coat collapse upon hydrolysis.

## MATERIALS AND METHODS

### Protein preparation for structural studies

Recombinant *M. musculus* coatomer was expressed and purified from SF9 insect cells (Invitrogen, Karlsruhe). Original SF9 cells were cloned from the parental IPLBSF-21 (Sf-21) cell line that was derived from the pupal ovarian tissue of the fall army worm, *Spodoptera frugiperda*. Invitrogen Sf9 cells were tested by the manufacturer for contamination of bacteria, yeast, mycoplasma and virus and were characterized by isozyme and karyotype analysis. We used a baculoviral expression system essentially as described previously (Sahlmuller et al., 2011) with a "One-Strep-Tag" at the C-terminus of  $\alpha$ -COP. Recombinant *S. cerevisiae* myristoylated Arf1, and human nucleotide exchange factor ARNO, were purified from *E. coli* as described previously (Chardin et al., 1996; Randazzo et al., 1992). Full-length N-terminally His-tagged *R. norvegicus* ArfGAP2 and ArfGAP1 proteins were expressed in insect cells and purified through Ni-NTA chromatography (Weimer et al., 2008), and gel filtration.

To determine the most stable domains of coatomer suitable for crystallization, limited proteolysis with subtilisin (Sigma Aldrich, St. Louis Missouri USA) was performed with *C. thermophilum* (Ct) coatomer and subcomplexes. 45-250  $\mu$ g of the respective complex was treated at various molar ratios. The reaction was incubated for 15 minutes on ice and then stopped by addition of PMSF to a final concentration of 1 mM. The samples were separated by SDS-PAGE and the resulting fragments were analyzed by MS.

Expression plasmids for truncated forms of Ct $\beta$  $\delta$ -COP subcomplexes were constructed using the pFBDM vector with a One-Strep-Tag fused the N-terminus of  $\beta$ -COP (Berger et al., 2004; Fitzgerald et al., 2006). Baculoviruses were generated for the subcomplexes Ct $\beta$ 19-391 $\delta$ , Ct $\beta$ 19-391 $\delta$ 1-243, Ct $\beta$ 19-391 $\delta$ 1-175 and Ct $\beta$ 19-391 $\delta$ 1-137 by infecting

Sf9 cells (Invitrogen, Karlsruhe) with recombinant Bacmids prepared using *E.coli* Dh10b MultiBac cells. The dimeric coatomer subcomplexes were produced by co-expression of both subunits in Sf9 insect cells infected with the corresponding baculovirus. Insect cells were harvested 72 h post infection. Cells were lysed in buffer (25 mM Tris, 300 mM NaCl, 1 mM DTT pH 8.0) using a high pressure Microfluidizer (Microfluidics, Newton USA) and cell debris was pelleted by centrifugation at 100 000 x g for 1 hr. The protein complex was purified using strep-tactin affinity chromatography according to the supplier's instructions, followed by size exclusion chromatography (Superdex 75 column, GE Healthcare). Purified protein was concentrated using Amicon spin concentrators (Merck Millipore, Darmstadt) and stored at -80°C.

In order to obtain phases by anomalous x-ray diffraction Ct $\beta$ 19-391 $\delta$ 1-175 subcomplex incorporating selenomethionine (MSE) was produced in Sf9 insect cells. Sf9 cells were grown in  $\Delta$ 921 Series methionine deficient medium (Expression Systems LLC, USA) supplemented with 150 mg/ml MSE.

## **Crystallization, X-ray Structure Determination, Analysis and Representation of $\beta\delta$ -COP**

Crystallization trials were performed with all construct variants described above and diffracting crystals could be obtained with polyethylene glycols as precipitant with a size range from PEG 3350 to PEG 8000. Best diffracting crystals were obtained with construct Ct $\beta$ 19-391 $\delta$ 1-175 using sitting drop vapor diffusion in 96-well MRC UVP plates at 18°C with a precipitant and reservoir composition of 0.2 M magnesium formate, 0.1 M Tris pH 7.0, and 24 % PEG3350. Crystallization drops consisted of 400 nl protein solution of Ct $\beta$ 19-391 $\delta$ 1-175 at a concentration of 13 mg/ml and 400 nl of

precipitant. The reservoir volume was 95  $\mu$ l. Crystals were visible after 10 days. All measured crystals had the orthorhombic space group  $C222_1$  with typical cell dimensions of  $a=137.59$ ,  $b=177.47$ ,  $c=62.72$ ,  $\alpha=\beta=\gamma=90^\circ$ .

All X-ray diffraction data were collected at ESRF beamlines ID-29 (native) and ID-21 (MSE). Anomalous diffraction data was collected from crystals containing MSE labelled Ct $\beta$ 19-391 $\delta$ 1-175. All data were integrated using XDS (Kabsch, 2010) and scaled with AIMLESS (Evans, 2006). Experimental phases were obtained with Phenix. autosol (Terwilliger et al., 2009) from anomalous diffraction data and an initial model was obtained using Phenix. autobuild (Terwilliger et al., 2008). The final model was built with Coot (Emsley et al., 2010) and refined with Phenix.refine (Afonine et al., 2012) using the native dataset (see Table 1). The asymmetric unit contains one molecule of Ct $\beta$ 20-390 $\delta$ 1-150. Structure quality was evaluated using MolProbity (Chen et al., 2010). Interfaces and crystal contacts were analyzed using PISA (Krissinel and Henrick, 2007). Structure figures were created with Chimera (Pettersen et al., 2004).

## **Pulldown assays**

For pulldown assays Arf1 (N $\Delta$ 20CtArf) was cloned in the petM11 vector, adding an N-terminal HIS<sub>6</sub>-TAG, and expressed in *E. coli* BL21 (DE3) cells (Novagen, Merck, Darmstadt). N $\Delta$ 20CtArf was purified by Ni-NTA affinity chromatography followed by gel filtration on a Superdex 75 column (GE Healthcare, Buckinghamshire UK) in PD (Pulldown) buffer (25 mM Tris, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT, pH 7.5).

To test coatamer  $\beta\delta$  subcomplexes for binding of Arf1, 200 to 400  $\mu$ g of the respective subcomplex was immobilized on 50  $\mu$ l strep-tactin sepharose beads by incubation for one hour at 4  $^\circ$ C on a rotary wheel. After immobilization of the subcomplex the beads

were washed with PD buffer to remove unbound protein. The soluble form of Arf1-GTP was added to the immobilized subcomplexes and incubated at 4 °C for one hour on a rotary wheel. After the incubation the beads were washed extensively with PD buffer. Next the beads were pelleted by centrifugation, resuspended in 50 µl buffer and samples were taken for each subcomplex and mixed with SDS sample buffer. After centrifugation, the proteins in the supernatants were analyzed by SDS-PAGE and western blot.

### **Site directed UV-Cross-linking**

For photo-cross-linking experiments human Arf1 and yeast N-myristoyl-transferase were cloned in pETDuet vector. Amber stop codons were introduced in Arf1 at positions 46 and 167 by point mutation. Methionine aminopeptidase for improvement of myristoylation efficiency was cloned into pRSF-Duet1 vector. Both plasmids were co-transformed in *E. coli* Bl21 (DE3) cells harbouring the pEVOL plasmid coding for orthogonal tRNA recognizing the Amber stop codon and the tRNA synthetase (Chin et al., 2002). Cells were grown to OD<sub>600</sub> of 0.6 at 37°C. After addition of 65 µM of sodium myristate and 1 mM of p-benzoyl-L-phenylalanine, cells were shifted to 27°C. Expression was induced 1 hr after the temperature shift by addition of IPTG (0.5 mM) and arabinose (0.5 %) and was continued for 22 hrs at 27°C. Photolabile Arf-I46Bp-Y167Bp was purified from the cleared lysate in the presence of 2 mM GDP by size exclusion chromatography on Superdex 200 (GE Healthcare, Buckinghamshire UK) in 25 mM Tris, 150 mM KCl, pH 7.0.

This bivalently photolabile Arf1 was used in a reconstitution assay with liposomes, GTP and complete COPI. For the reaction 5 µg Arf-I46Bp-Y167Bp was mixed with 0.4 µg

ARNO, 100  $\mu$ M GTP $\gamma$ S and 500  $\mu$ M Golgi-like liposomes in a final volume of 100  $\mu$ l in 20 mM MOPS pH 7.2, 150 mM KOAc 2 mM Mg(OAc)<sub>2</sub>, and incubated at 37 °C for 10 minutes. Alternatively 50  $\mu$ g of Golgi membranes in the presence of 200 mM Sucrose were incubated with 5  $\mu$ g of the respective photolabile Arf1 with or without 100  $\mu$ M GTP $\gamma$ S in a final volume of 100  $\mu$ l buffer at 37 °C for 10 minutes. In a second step, coatomer was added and incubation was resumed at 37°C for 10 minutes. Membranes were pelleted by centrifugation at 16,000 x g for 30 minutes at 4 °C. Golgi membranes from the liver of male 200 g Wistar rats (*Rattus norvegicus*) were pelleted through 330  $\mu$ l of 15% (v/v) Sucrose. The pellet was re-suspended in 10  $\mu$ l buffer and irradiated on ice with 15 x 1 second UV<sub>366</sub> pulses with 1 second pauses. After irradiation the sample was analysed by MS, SDS-PAGE, and western blot using antibodies directed against  $\beta$ -COP,  $\delta$ -COP and Arf1. Additionally photolabile amino acids were introduced in position 156 or 159 of  $\delta$ -COP (*M. musculus*) using the same approach. These  $\delta$ -COP variants were introduced in  $\beta\delta$ -COP subcomplexes together with Ct $\beta$ 19-391 producing a chimeric  $\beta\delta$  subcomplex. This was done as no antibodies were available against Ct $\delta$ -COP. The chimeric subcomplexes were tested in the reconstitution assay with liposomes, GTP and Arf1 as described above.

#### **In vitro budding reaction**

Giant Unilamellar Vesicles (GUVs) were prepared by electroformation (Angelova et al., 1992) from the Golgi-like lipid mix (Bigay and Antonny, 2005). COPI-coated vesicles were produced in vitro by incubating coatomer (840 nM), Arf1 (2  $\mu$ M), GTP $\gamma$ S (1 mM), ARNO (1.5  $\mu$ M) and 2  $\mu$ l GUVs in a total volume of 40  $\mu$ l for 30 minutes at 37°C. The budding reaction buffer contained 50 mM HEPES pH 7.4, 50 mM KOAc, 1 mM MgCl<sub>2</sub>.

Protein-A conjugated 10 nm gold was added to the reaction mix in 1:6 volume ratio and the sample was applied onto glow-discharged (30 sec, 20 mA) C-flat (Protochips Inc.) multihole grids. The grids were blotted from the back side for 11 seconds at room temperature in a chamber at 85% humidity and plunge-frozen into liquid ethane using a manual plunger.

In order to test activity of the ArfGAP1 and ArfGAP2 proteins, the COPI budding reaction was performed in the presence of GTP, and the reaction mix was incubated for 30 minutes at 37°C. ArfGAP1, ArfGAP2 or buffer (as a control) were added to the mix in 10 molar excess to coatomer and after 15 minutes incubation the reaction was plunge-frozen. All samples were imaged in an electron microscope. The control samples contained coated vesicles and buds, whereas the samples incubated with ArfGAP1 or ArfGAP2 contained only naked liposomes. The functionality of ArfGAP proteins in vitro was also shown previously in budding assays containing rat liver Golgi membranes (Weimer et al., 2008).

To explore the structure of the coat in the presence of ArfGAP1 or ArfGAP2, the budding reaction mix was incubated for 30 minutes prior to addition of a 10 molar excess of ArfGAP1 or ArfGAP2 in the presence of GTPγS nucleotide. The mix was incubated for a further 15-20 minutes, protein-A conjugated 10 nm gold was added in a 1:6 volume ratio, and the sample was plunge-frozen.

#### **CryoET sample preparation, data acquisition and initial processing**

The plunge-frozen COPI-coated vesicles were imaged in a FEI Titan Krios electron microscope operated at 300 kV and equipped with a Gatan Quantum 967 LS energy filter with a 20 eV energy slit and Gatan K2xp direct electron detector (Gatan Inc.).

Tomographic tilt series were acquired with the dose-symmetric tilt-scheme (Hagen et al., 2016) over an angular range of  $\pm 60^\circ$  with a  $3^\circ$  increment and a total electron dose of approximately  $85 \text{ e}/\text{\AA}^2$ . The defocus values ranged from  $-2.0$  to  $-5.0 \text{ }\mu\text{m}$ . Data acquisition was controlled using the SerialEM software package (Mastronarde, 2005). Five frames were collected in super-resolution and electron-counting mode at each tilt. The super-resolution pixel size at the specimen level was  $0.89 \text{ }\text{\AA}$ . The frames were Fourier-cropped, motion-corrected with the K2Align package based on the MotionCorr algorithms (Li et al., 2013) and integrated together. Each of the images in the tilt series was low-pass filtered according to the electron-dose acquired by the sample (Grant and Grigorieff, 2015). Motion-corrected and dose-filtered tomograms were reconstructed in Imod (Kremer et al., 1996).

## **Image processing**

1733 vesicles and near-complete buds were picked from 61 tomograms. CTF-determination for each individual tilt image was performed using CTFFIND4 (Rohou and Grigorieff, 2015). Strip-based CTF-correction and tomogram reconstruction was performed in Imod. Subtomograms were extracted from the surface of the vesicles. Subtomogram averaging was performed using scripts derived from the TOM and Av3 software packages (Forster et al., 2005; Nickell et al., 2005) and Dynamo (Castano-Diez et al., 2012). Each dataset was split into two halves, each half including odd- and even-numbered vesicles. The COPI triad structure (EMDB-2985, (Dodonova et al., 2015)) was low pass filtered to  $55 \text{ }\text{\AA}$  and used for the initial alignment step. All further processing steps were performed completely independently on the two half datasets. Preliminary processing was performed using 4x binned data without CTF correction (pixel size  $7.12$



Å). C3 symmetry was applied. The references were low pass filtered to 35 Å at each alignment step. Iterative rotational and translational alignments were performed until convergence. Next the dataset was cleaned to remove misaligned subtomograms based on cross-correlation coefficient threshold, and duplicates were removed based on the mutual distance between neighboring subtomograms. Final subtomogram alignments were performed on the unbinned (pixel size 1.78 Å) and CTF-corrected data with the low pass filter set to 16 Å. The half datasets contained 19,343 and 19,155 asymmetric units. At the final alignment steps each asymmetric unit of the triad, “the leaf”, representing a COPI molecule and two Arf1 molecules, was processed separately in order to compensate for coat movements and deviation from C3 symmetry. The FSC was calculated for the two final references masked with a soft cylindrical mask and the measured resolution at FSC=0.143 was 9.4 Å (Figure 2 – Figure Supplement 1 A).

The local resolution of the EM map was estimated within a small floating window moving within the whole map. The local resolution was highly variable and ranged from 8.7 to 12.7 Å within the protein density (Figure 2 – Figure Supplement 1 B). We therefore performed multiple local alignments using soft cylindrical masks focused on different parts of the structure. Finally, all the local averages were masked with the corresponding alignment masks, added together, and the map was normalized within the volume enclosed by all local masks. The local resolution ranged from 8.1 to 11.5 Å. The global resolution was 9.2 Å at FSC 0.143. The final structure (Figure 2B) was B-factor sharpened (B-factor=-1400) (Rosenthal and Henderson, 2003).

## **Structural determination of COPI linkages**

The COPI triads can be arranged into four distinct types of linkages within the coat (Faini et al., 2012). We used the approach described in detail in (Dodonova et al., 2015; Faini et al., 2012) to determine their positions and orientations. Next we performed subtomogram averaging on each of the linkage datasets to determine their structures (Figure 2 – Figure Supplement 2). Each set was processed in two independent halves, which were split based on the odd or even sequential number of the vesicle. C2 symmetry was applied to the linkage III and IV sets. The starting references were obtained by averaging the subvolumes at their initial positions. The alignments were performed iteratively starting with a full search for the in-plane Euler angle and continued with a progressively decreasing angular search step. Duplicate and misaligned subtomograms were removed. Next the subvolumes were extracted at the positions defined during previous iterations from the CTF-corrected unbinned tomograms, and final alignments were performed. The linkage datasets contained 2547, 3312, 140, 1640 asymmetric units from both half sets. The final resolutions of linkages I, II, and IV at FSC 0.143 were respectively 17, 15, 17.3 Å. Linkage III had low abundance in the dataset and the resulting resolution was 31 Å.

In order to generate pseudo-atomic models of the linkages, first, the model of the COPI leaf generated by homology modeling and flexible fitting into the 9.2 Å leaf EM map (see below) was fitted as a rigid body into each of the linkage EM maps. Next, homology models of linkage-specific subunits ( $\alpha$ -COP C-terminal domain together with  $\epsilon$ -COP and the MHD of  $\delta$ -COP) were generated in Modeller and fitted as rigid bodies into the central densities in the linkage EM maps.

#### **Homology modeling and fitting**

First the available crystal structures of COPI components were fitted as rigid bodies into the structure of the leaf: Arf1- $\gamma$ -COP (PDB:3TJZ (Yu et al., 2012)),  $\alpha\beta'$ -COP (PDB:3MKQ (Lee and Goldberg, 2010)), the N-terminal  $\beta$ -propeller of  $\alpha$ -COP (PDB:4J87 (Ma and Goldberg, 2013)), and the  $\gamma$ -COP appendage (PDB:1PZD (Hoffman et al., 2003)). Moreover, several crystal structures corresponding to the densities in the inter-triad linkages were also available: the  $\alpha\epsilon$ -COP structure (PDB:3MKR and 3MV2 (Hsia and Hoelz, 2010; Lee and Goldberg, 2010)), the  $\delta$ -COP MHD (PDB:4O8Q (Lahav et al., 2015)).

Homology models of COPI subunits were generated automatically using the HHpred and Modeller servers (Soding et al., 2005; Webb and Sali, 2014), as described previously (Dodonova et al., 2015). The model for  $\beta$ -COP (residues 410-968) was generated based on the structures of the homologous AP2  $\beta$ 2 subunit (PDB:2XA7 and PDB:2G30) (Edeling et al., 2006; Jackson et al., 2010); the model for  $\gamma$ -COP (residues 312-549) was generated based on the structure of the AP2  $\alpha$ 2 (PDB:2XA7) (Jackson et al., 2010).  $\alpha$ -COP residues 327-813 were modeled based on the  $\beta'$ -COP structure (PDB:3MKQ) (Lee and Goldberg, 2010). We performed “structure-guided” modeling for the C-terminal parts of the  $\beta$ -COP and  $\gamma$ -COP trunk domains, which were absent from the starting x-ray structural models. The  $\gamma$ -COP and  $\beta$ -COP trunk C-terminal parts were modeled as separate batches of 2-3 helices, which were sequentially fitted into the distinct  $\alpha$ -helical EM densities. Loops connecting the batches of helices were added in Modeller.

The structures of all remaining COPI subunits or parts of subunits were available from the PDB and were modeled to represent the *M. musculus* sequence.

Rigid body fitting was performed in Chimera (Pettersen et al., 2004). Flexible fitting was performed using NAMD with the MDFF package (Trabuco et al., 2008).

691

## 692 **Extra density assignment**

693 After flexible fitting of all COPI subunits into the map, several small regions of the map  
694 remained unoccupied. The unoccupied “extra” densities were identified by subtracting  
695 the density occupied by the fitted model from the EM map. The density occupied by the  
696 fitted model was generated by the chimera molmap command at 9 Å resolution. The  
697 resulting difference map was segmented and the volumes of all separate extra densities  
698 were plotted (Figure 4 – Figure Supplement 1). The seven large extra densities were  
699 numbered according to their sizes (Figure 4 – Figure Supplement 1, numbered): “1”  
700 near the  $\beta'$ - $\alpha$ -COP interface; “2” near the N-terminal  $\beta$ -propeller domain of  $\alpha$ -COP; “3”  
701 and “4” near the  $\beta$ - $\gamma$ -COP trunks interface; “5” near a  $\beta$ - $\beta'$ -COP contact site; “6” near the  
702  $\beta$ - $\delta$  interface; “7” near the Arf1- $\beta$ - $\delta$  interface;

703

## 704 **COPI-ArfGAP2 dataset**

705 The COPI-ArfGAP2 dataset consisted of 27 tomograms containing 690 vesicles and  
706 near-complete buds. Subtomogram averaging and alignments were performed for two  
707 independent half datasets, which contained either odd- or even-numbered vesicles and  
708 consisted of 11,331 and 11,004 asymmetric units respectively. The subtomogram  
709 averaging processing pipeline was exactly the same as for the COPI dataset. The final  
710 measured resolution of the leaf at the FSC=0.143 was 9.8 Å. We calculated the difference  
711 map between the final COPI-ArfGAP2 EM map and the control COPI EM map. An  
712 additional density was visible near the Arf1 molecules located at the center of the triad.  
713 Multireference alignment and classification on the complete leaf structure was  
714 performed in order to identify the subpopulation of leaves with bound ArfGAP2. To do

this, the subtomograms were aligned against the reference from the COPI-ArfGAP2-dataset and from the COPI-dataset and divided into two classes based on cross-correlation. The average structures were generated for both classes. This process was iterated a total of three times. The final COPI-ArfGAP2 class contained 6280 and 6092 asymmetric units in two independent half datasets, comprising in total approximately 65 % of the initial dataset. The resolution was 10.1 Å at FSC=0.143. The second class, which did not contain additional ArfGAP density, consisted of 3497 and 3481 asymmetric units and the resolution at the 0.143 FSC was measured to be 11.7 Å. The final difference map was calculated between the structures produced from the two classes (Figure 7 – Figure Supplement 1).

The linkage maps from the COPI-ArfGAP2 were calculated and upon comparison with the control maps did not show any additional densities except that described above.

#### **COPI-ArfGAP1 dataset**

The COPI-ArfGAP1 dataset comprised 24 tomograms containing 833 vesicles and near-complete buds. Subtomogram averaging and alignments were performed for two independent halves of the dataset. The two half datasets contained 14,613 and 14,379 asymmetric units. The subtomogram averaging processing pipeline was exactly the same as for the COPI and COPI-ArfGAP2 datasets. The final measured resolution of the COPI leaf structure at the FSC=0.143 was 9.6 Å. We calculated the difference map between the final COPI-ArfGAP1 leaf EM map and the control COPI EM map. No large additional densities were observed. The structures of the linkages were calculated and upon comparison with the control maps did not show any additional densities.

739

740 **ACKNOWLEDGMENTS**

741 EM maps and corresponding structural models are deposited in the Electron  
742 Microscopy Data Bank (accession codes EMD-3720, EMD-3721, EMD-3722, EMD-3723,  
743 EMD-3724) and the Protein Data Bank (accession codes PDB: 5NZR, 5NZS, 5NZT, 5NZU,  
744 5NZV). Coordinates and structure factors for the X-ray structure of  $\beta\delta$ -COP are  
745 deposited in PDB under accession code 5MU7. We thank P. Schultz for the vectors used  
746 for site directed photolabeling, K. Bacia for providing *S. cerevisiae* Arf1 protein; A. von  
747 Appen and J. Kosinski for help with the MS data; F. Schur, W. Wan, O. Avinoam, S. Mattei,  
748 Y. Bykov for discussions. We are grateful to J. Goldberg for providing the coordinates of  
749 the Arf1-ArfGAP1 model. We thank C. Siegmann from the BZH/Cluster of  
750 Excellence:CellNetworks crystallization platform for support in protein crystallization.  
751 We thank the staff of ESRF and of EMBL-Grenoble for assistance and support in using  
752 beamlines ID-21 and ID-29. Our work was technically supported by EMBL IT services,  
753 and was funded by the Deutsche Forschungsgemeinschaft within SFB638 (A16) to JAGB  
754 and FW, and SFB638 (Z4) to IS; and WI 654/12-1 to FW. FW and IS are investigators of  
755 the Cluster of Excellence:CellNetworks.  
756 Hereby we declare no financial and non-financial competing interests.

757

## 758 REFERENCES

- 759 Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W.,  
760 Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012).  
761 Towards automated crystallographic structure refinement with phenix.refine. *Acta*  
762 *Crystallogr D Biol Crystallogr* **68**, 352-367.
- 763 Alisaraie, L., and Rouiller, I. (2012). Full-length structural model of RET3 and SEC21 in  
764 COPI. *J Mol Model* **18**, 3199-3212.
- 765 Amor, J.C., Harrison, D.H., Kahn, R.A., and Ringe, D. (1994). Structure of the human ADP-  
766 ribosylation factor 1 complexed with GDP. *Nature* **372**, 704-708.
- 767 Angelova, M., Soleau, S., Meleard, P., Faucon, J., and Bothorel, P. (1992). Preparation of  
768 giant vesicles by external AC electric fields. *Progress in Colloid & Polymer Science* **89**,  
769 127-131.
- 770 Arakel, E.C., Richter, K.P., Clancy, A., and Schwappach, B. (2016). Delta-COP contains a  
771 helix C-terminal to its longin domain key to COPI dynamics and function. *Proc Natl Acad*  
772 *Sci U S A* **113**, 6916-6921.
- 773 Berger, I., Fitzgerald, D.J., and Richmond, T.J. (2004). Baculovirus expression system for  
774 heterologous multiprotein complexes. *Nat Biotechnol* **22**, 1583-1587.
- 775 Bigay, J., and Antonny, B. (2005). Real-time assays for the assembly-disassembly cycle of  
776 COP coats on liposomes of defined size. *Methods Enzymol* **404**, 95-107.
- 777 Bigay, J., Casella, J.F., Drin, G., Mesmin, B., and Antonny, B. (2005). ArfGAP1 responds to  
778 membrane curvature through the folding of a lipid packing sensor motif. *EMBO J* **24**,  
779 2244-2253.
- 780 Castano-Diez, D., Kudryashev, M., Arheit, M., and Stahlberg, H. (2012). Dynamo: a  
781 flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in  
782 high-performance computing environments. *J Struct Biol* **178**, 139-151.
- 783 Chardin, P., Paris, S., Antonny, B., Robineau, S., Beraud-Dufour, S., and Chabre, M. (1996).  
784 A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains.  
785 *Nature* **384**, 481-484.
- 786 Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J.,  
787 Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom  
788 structure validation for macromolecular crystallography. *Acta Crystallogr D Biol*  
789 *Crystallogr* **66**, 12-21.
- 790 Chin, J.W., Martin, A.B., King, D.S., Wang, L., and Schultz, P.G. (2002). Addition of a  
791 photocrosslinking amino acid to the genetic code of *Escherichiacoli*. *Proc Natl Acad Sci U*  
792 *S A* **99**, 11020-11024.
- 793 Cosson, P., and Letourneur, F. (1994). Coatamer interaction with di-lysine endoplasmic  
794 reticulum retention motifs. *Science* **263**, 1629-1631.
- 795 Cukierman, E., Huber, I., Rotman, M., and Cassel, D. (1995). The ARF1 GTPase-activating  
796 protein: zinc finger motif and Golgi complex localization. *Science* **270**, 1999-2002.
- 797 DeRegis, C.J., Rahl, P.B., Hoffman, G.R., Cerione, R.A., and Collins, R.N. (2008). Mutational  
798 analysis of betaCOP (Sec26p) identifies an appendage domain critical for function. *BMC*  
799 *Cell Biol* **9**, 3.
- 800 Devos, D., Dokudovskaya, S., Alber, F., Williams, R., Chait, B.T., Sali, A., and Rout, M.P.  
801 (2004). Components of coated vesicles and nuclear pore complexes share a common  
802 molecular architecture. *PLoS Biol* **2**, 2085-2093.
- 803 Dodonova, S.O., Diestelkoetter-Bachert, P., von Appen, A., Hagen, W.J., Beck, R., Beck, M.,  
804 Wieland, F., and Briggs, J.A. (2015). A structure of the COPI coat and the role of coat  
805 proteins in membrane vesicle assembly. *Science* **349**, 195-198.

Edeling, M.A., Mishra, S.K., Keyel, P.A., Steinhäuser, A.L., Collins, B.M., Roth, R., Heuser, J.E., Owen, D.J., and Traub, L.M. (2006). Molecular switches involving the AP-2  $\beta$ 2 appendage regulate endocytic cargo selection and clathrin coat assembly. *Dev Cell* 10, 329-342.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501.

Eugster, A., Frigerio, G., Dale, M., and Duden, R. (2000). COP I domains required for coatamer integrity, and novel interactions with ARF and ARF-GAP. *The EMBO journal* 19, 3905-3917.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62, 72-82.

Faini, M., Prinz, S., Beck, R., Schorb, M., Riches, J.D., Bacia, K., Brugger, B., Wieland, F.T., and Briggs, J.A. (2012). The structures of COPI-coated vesicles reveal alternate coatamer conformations and interactions. *Science* 336, 1451-1454.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., and Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. *Nat Methods* 3, 1021-1032.

Forster, F., Medalia, O., Zauberman, N., Baumeister, W., and Fass, D. (2005). Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc Natl Acad Sci U S A* 102, 4729-4734.

Fotin, A., Cheng, Y., Sliz, P., Grigorieff, N., Harrison, S.C., Kirchhausen, T., and Walz, T. (2004). Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* 432, 573-579.

Frigerio, G., Grimsey, N., Dale, M., Majoul, I., and Duden, R. (2007). Two human ARFGAPs associated with COP-I-coated vesicles. *Traffic* 8, 1644-1655.

Goldberg, J. (1998). Structural basis for activation of ARF GTPase: Mechanisms of Guanine Nucleotide Exchange and GTP-Myristoyl Switching. *Cell* 95, 237-248.

Goldberg, J. (1999). Structural and functional analysis of the ARF1-ARFGAP Complex Reveals a Role for Coatamer in GTP Hydrolysis. *Cell* 96, 893-902.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife* 4.

Hagen, W.J., Wan, W., and Briggs, J.A. (2016). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J Struct Biol*.

Hara-Kuge, S., Kuge, O., Orci, L., Amherdt, M., Ravazzola, M., Wieland, F., and Rothman, J. (1994). En bloc incorporation of coatamer subunits during the assembly of COP-coated vesicles. *J Cell Biol* 124, 883-892.

Harter, C., and Wieland, F. (1998). A single binding site for dilysine retrieval motifs and p23 within the  $\gamma$  subunit of coatamer. *Proc Natl Acad Sci USA* 95, 11649-11654.

Hoffman, G.R., Rahl, P.B., Collins, R.N., and Cerione, R.A. (2003). Conserved structural motifs in intracellular trafficking pathways: structure of the gammaCOP appendage domain. *Mol Cell* 12, 615-625.

Hsia, K.-C., and Hoelz, A. (2010). Crystal structure of  $\alpha$ -COP in complex with  $\epsilon$ -COP provides insight into the architecture of the COPI vesicular coat. *Proc Natl Acad Sci USA* 107, 11271-11276.

Ismail, S.A., Vetter, I.R., Sot, B., and Wittinghofer, A. (2010). The structure of an Arf-ArfGAP complex reveals a  $\text{Ca}^{2+}$  regulatory mechanism. *Cell* 141, 812-821.



Jackson, L.P., Kelly, B.T., McCoy, A.J., Gaffry, T., James, L.C., Collins, B.M., Honing, S., Evans, P.R., and Owen, D.J. (2010). A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. *Cell* **141**, 1220-1229.  
 Jackson, L.P., Lewis, M., Kent, H.M., Edeling, M.A., Evans, P.R., Duden, R., and Owen, D.J. (2012). Molecular Basis for Recognition of Dilysine Trafficking Motifs by COPI. *Dev Cell*.  
 Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.  
 Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132.  
 Kimata, Y., Higashio, H., and Kohno, K. (2000). Impaired proteasome function rescues thermosensitivity of yeast cells lacking the coatamer subunit epsilon-COP. *J Biol Chem* **275**, 10655-10660.  
 Kliouchnikov, L., Bigay, J., Mesmin, B., Parnis, A., Rawet, M., Goldfeder, N., Antonny, B., and Cassel, D. (2009). Discrete determinants in ArfGAP2/3 conferring Golgi localization and regulation by the COPI coat. *Mol Biol Cell* **20**, 859-869.  
 Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76.  
 Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797.  
 Lahav, A., Rozenberg, H., Parnis, A., Cassel, D., and Adir, N. (2015). Structure of the bovine COPI delta subunit mu homology domain at 2.15 Å resolution. *Acta Crystallogr D Biol Crystallogr* **71**, 1328-1334.  
 Lee, C., and Goldberg, J. (2010). Structure of coatamer cage proteins and the relationship among COPI, COPII, and clathrin vesicle coats. *Cell* **142**, 123-132.  
 Letourneur, F., Gaynor, E.C., Hennecke, S., Démollière, C., Duden, R., Emr, S.D., Riezman, H., and Cosson, P. (1994). Coatamer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* **79**, 1199-1207.  
 Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584-590.  
 Liu, W., Duden, R., Phair, R.D., and Lippincott-Schwartz, J. (2005). ArfGAP1 dynamics and its role in COPI coat assembly on Golgi membranes of living cells. *J Cell Biol* **168**, 1053-1063.  
 Lowe, M., and Kreis, T.E. (1995). In vitro assembly and disassembly of coatamer. *J Biol Chem* **270**, 31364-31371.  
 Luo, R., Ha, V.L., Hayashi, R., and Randazzo, P.A. (2009). Arf GAP2 is positively regulated by coatamer and cargo. *Cell Signal* **21**, 1169-1179.  
 Ma, W., and Goldberg, J. (2013). Rules for the recognition of dilysine retrieval motifs by coatamer. *EMBO J*.  
 Majoul, I., Straub, M., Hell, S.W., Duden, R., and Soling, H.D. (2001). KDEL-cargo regulates interactions between proteins involved in COPI vesicle traffic: Measurements in living cells using FRET. *Developmental cell* **1**, 139-153.  
 Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51.  
 Nickell, S., Forster, F., Linaroudis, A., Net, W.D., Beck, F., Hegerl, R., Baumeister, W., and Plitzko, J.M. (2005). TOM software toolbox: acquisition and analysis for electron tomography. *J Struct Biol* **149**, 227-234.  
 Orci, L., Stamnes, M., Ravazzola, M., Amherdt, M., Perrelet, A., Sollner, T.H., and Rothman, J.E. (1997). Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell* **90**, 335-349.

902 Ouali, M., and King, R.D. (2000). Cascaded multiple classifiers for secondary structure  
903 prediction. *Protein Sci* 9, 1162-1176.

904 Owen, D.J., Vallis, Y., Pearse, B.M., McMahon, H.T., and Evans, P.R. (2000). The structure  
905 and function of the beta 2-adaptin appendage domain. *EMBO J* 19, 4216-4227.

906 Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and  
907 Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and  
908 analysis. *J Comput Chem* 25, 1605-1612.

909 Pevzner, I., Strating, J., Lifshitz, L., Parnis, A., Glaser, F., Herrmann, A., Brugger, B.,  
910 Wieland, F., and Cassel, D. (2012). Distinct role of subcomplexes of the COPI coat in the  
911 regulation of ArfGAP2 activity. *Traffic* 13, 849-856.

912 Pietroseoli, N., Pancsa, R., and Tompa, P. (2013). Structural disorder provides  
913 increased adaptability for vesicle trafficking pathways. *PLoS Comput Biol* 9, e1003144.

914 Poon, P.P., Cassel, D., Spang, A., Rotman, M., Pick, E., Singer, R.A., and Johnston, G.C.  
915 (1999). Retrograde transport from the yeast Golgi is mediated by two ARF GAP proteins  
916 with overlapping function. *EMBO J* 18, 555-564.

917 Praefcke, G.J., Ford, M.G., Schmid, E.M., Olesen, L.E., Gallop, J.L., Peak-Chew, S.Y., Vallis, Y.,  
918 Babu, M.M., Mills, I.G., and McMahon, H.T. (2004). Evolving nature of the AP2 alpha-  
919 appendage hub during clathrin-coated vesicle endocytosis. *EMBO J* 23, 4371-4383.

920 Randazzo, P.A., Weiss, O., and Kahn, R.A. (1992). Preparation of recombinant ADP-  
921 ribosylation factor. *Methods Enzymol* 219, 362-369.

922 Ren, X., Farias, G.G., Canagarajah, B.J., Bonifacino, J.S., and Hurley, J.H. (2013). Structural  
923 basis for recruitment and activation of the AP-1 clathrin adaptor complex by Arf1. *Cell*  
924 152, 755-767.

925 Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation  
926 from electron micrographs. *J Struct Biol* 192, 216-221.

927 Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle  
928 orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy.  
929 *J Mol Biol* 333, 721-745.

930 Sahlmuller, M.C., Strating, J.R., Beck, R., Eckert, P., Popoff, V., Haag, M., Hellwig, A., Berger,  
931 I., Brugger, B., and Wieland, F.T. (2011). Recombinant heptameric coatomer complexes:  
932 novel tools to study isoform-specific functions. *Traffic* 12, 682-692.

933 Saitoh, A., Shin, H.W., Yamada, A., Waguri, S., and Nakayama, K. (2009). Three  
934 homologous ArfGAPs participate in coat protein I-mediated transport. *J Biol Chem* 284,  
935 13948-13957.

936 Schledzewski, K., Brinkmann, H., and Mendel, R.R. (1999). Phylogenetic analysis of  
937 components of the eukaryotic vesicle transport system reveals a common origin of  
938 adaptor protein complexes 1,2 and 3 and the F subcomplex of the coatomer COPI. *J Mol*  
939 *Evol* 48, 770-778.

940 Schmid, E.M., Ford, M.G., Burtsey, A., Praefcke, G.J., Peak-Chew, S.Y., Mills, I.G., Benmerah,  
941 A., and McMahon, H.T. (2006). Role of the AP2  $\beta$ -appendage hub in recruiting partners  
942 for clathrin-coated vesicle assembly. *PLoS Biol* 4, e262.

943 Semenza, J.C., Hardwick, K.G., Dean, N., and Pelham, H.R. (1990). ERD2, a yeast gene  
944 required for the receptor-mediated retrieval of luminal ER proteins from the secretory  
945 pathway. *Cell* 61, 1349-1357.

946 Shen, Q.T., Ren, X., Zhang, R., Lee, I.H., and Hurley, J.H. (2015). HIV-1 Nef hijacks clathrin  
947 coats by stabilizing AP-1:Arf1 polygons. *Science* 350, aac5137.

948 Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein  
949 homology detection and structure prediction. *Nucleic Acids Res* 33, W244-248.

Stagg, S.M., Gurkan, C., Fowler, D.M., LaPointe, P., Foss, T.R., Potter, C.S., Carragher, B., and Balch, W.E. (2006). Structure of the Sec13/31 COPII coat cage. *Nature* 439, 234-238.

Sun, Z., Anderl, F., Frohlich, K., Zhao, L., Hanke, S., Brugger, B., Wieland, F., and Bethune, J. (2007). Multiple and stepwise interactions between coatamer and ADP-ribosylation factor-1 (Arf1)-GTP. *Traffic* 8, 582-593.

Szafer, E., Rotman, M., and Cassel, D. (2001). Regulation of GTP hydrolysis on ADP-ribosylation factor-1 at the Golgi membrane. *J Biol Chem* 276, 47834-47839.

Terwilliger, T.C., Adams, P.D., Read, R.J., McCoy, A.J., Moriarty, N.W., Grosse-Kunstleve, R.W., Afonine, P.V., Zwart, P.H., and Hung, L.W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 65, 582-601.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64, 61-69.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16, 673-683.

Watson, P.J., Frigerio, G., Collins, B.M., Duden, R., and Owen, D.J. (2004).  $\gamma$ -COP Appendage Domain – Structure and Function. *Traffic* 5, 79-88.

Webb, B., and Sali, A. (2014). Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* 47.

Weimer, C., Beck, R., Eckert, P., Reckmann, I., Moelleken, J., Brugger, B., and Wieland, F. (2008). Differential roles of ArfGAP1, ArfGAP2, and ArfGAP3 in COPI trafficking. *J Cell Biol* 183, 725-735.

Yahara, N., Sato, K., and Nakano, A. (2006). The Arf1p GTPase-activating protein Glo3p executes its regulatory function through a conserved repeat motif at its C-terminus. *J Cell Sci* 119, 2604-2612.

Yu, X., Breitman, M., and Goldberg, J. (2012). A Structure-Based Mechanism for Arf1-Dependent Recruitment of Coatamer to Membranes. *Cell* 148, 530-542.

## FIGURE LEGENDS

### Figure 1

#### Crystal structure of $\beta$ 20-390 $\delta$ 2-150-COP

A ribbon representation of the structure of Ct $\beta$ 20-390 (dark green) Ct $\delta$ 2-150 (orange) as viewed from the vesicle membrane. The  $\alpha$ -helices within the  $\alpha$ -solenoid of  $\beta$ -COP are numbered according to their position in the sequence.  $\delta$ -COP consists of a longin domain and two downstream helices: helix a and b. Helix b projects away from the  $\alpha$ -solenoid of  $\beta$ -COP. The schematic at the bottom of the panel shows the subunit domain composition. The subunit regions that are visualised in the x-structure are marked. The  $\beta$ -COP sequence numbering includes the N-terminal His-tag and thus is shifted by 15 residues relative to the annotated uniprot-Q9JIF7 *Mus Musculus* sequence. See also Figure 1 – Figure Supplement 1 and Table 2.

### Figure 2

#### The structure of the COPI leaf at 9 Å resolution

(A) CryoET reconstruction of the COPI asymmetric unit, the “leaf” before local alignments. The density is colored according to the distance from the membrane - from red to blue. The displayed structure also contains part of Arf1- $\gamma$ - $\zeta$ -COP from a neighboring leaf to show inter-leaf interactions. The density is displayed at 0.04 isosurface level in order to visualize the membrane. (B) CryoET reconstruction of the leaf at 9.2 Å resolution after local alignment. The membrane was masked out during local alignments and upon generation of the combined map. Note the definition of the  $\alpha$ -helical densities in the structure. (C) A structural model of the COPI coat after flexible-fitting of structures and homology models into the cryoET structure. Note, that the C-

terminal domain of  $\alpha$ -COP,  $\varepsilon$ -COP, and the  $\delta$ -COP MHD, are not visualized in the leaf structure, since they compose the inter-triad linkages. Color scheme: cryoET density - grey, Arf1 - pink,  $\gamma$ -COP - light green,  $\beta$ -COP - dark green,  $\zeta$ -COP - yellow,  $\delta$ -COP - orange,  $\beta'$ -COP - light blue,  $\alpha$ -COP - dark blue. **(D)** The COPI triad. One asymmetric unit, the “leaf”, is outlined with an orange line. The part of the structure displayed in this figure A-C is outlined with the white line. The central Arf1 ( $\gamma$ Arf1) is marked with a blue asterisk, the peripheral Arf1 ( $\beta$ Arf1) with a red asterisk. COPI subunits are displayed as molecular surfaces. See also Figure 2 – Figure Supplement 1, Figure Supplement 2, and Video 1.

### **Figure 3**

#### **$\gamma$ -COP and $\beta$ -COP appendage domains within the coat**

**(A)** Localization of the  $\gamma$ -COP and  $\beta$ -COP appendages in the COPI leaf. The 9.2 Å EM map of the COPI leaf is colored based on the underlying subunit. Color scheme as in Figure 1, additionally appendage sandwich subdomains are red and appendage platform subdomains are purple. **(B)**  $\gamma$ -COP appendage (sandwich subdomain - red, platform subdomain - purple) interacts with the  $\beta'$ -COP (light blue). The models are shown within the corresponding part of the COPI EM map (transparent grey isosurface) **(C)** The structure as in B, opened out to reveal the conserved interaction interfaces of  $\gamma$ -COP and  $\beta'$ -COP (purple - conserved, cyan - variable). The yellow asterisks mark the interaction surfaces. **(D)**  $\beta$ -COP appendage interacts with  $\alpha$ -COP (dark blue). **(E)** The structure as in D, opened out to reveal the conserved interaction interfaces of  $\beta$ -COP appendage and  $\alpha$ -COP (purple - conserved, cyan - variable). **(F)** The  $\beta$ -COP appendage domain provides the main structural connection between the  $\gamma$ -COP adaptor subunit (light green) and the outer-coat subunit  $\alpha$ -COP (blue), where it interacts with both  $\beta$ -

propeller domains. Organization of the  $\beta$ -COP subunit: the long  $\beta$ -COP trunk domain (dark green) is connected to the appendage domain (red/purple) by a flexible linker (dashed orange line). We note that the  $\beta$ -COP trunk domain conformation, which is rather straight, is significantly different from the conformations of the  $\beta$ -subunits from homologous AP complexes, which are highly curved. The panel on the right shows the subcomplex from the “bottom” (from the membrane side) in order to visualize the suggested  $\beta$ -linker path more clearly.

#### **Figure 4**

##### **New densities identified in the COPI structure**

(A) Densities within the EM map that are not occupied by the fitted domain structures are shown in red and numbered from 1 to 7 according to their size. The fitted molecules in the lower panel are shown as surfaces. (B) The outer-coat subunits of COPI and a newly assigned region. Density 1 is adjacent to the  $\alpha$ -COP/ $\beta'$ -COP interface and can be fitted with the  $\alpha$ -helices (red) from the C-terminal part of the  $\beta'$ -COP subunit. (C) The density near the  $\alpha$ -COP N-terminal  $\beta$ -propeller domain. The extra density at least partly originates from the mobile loop on the side of the  $\beta$ -propeller domain. (D) The adaptor subunits of COPI and three newly assigned regions. Densities 3 and 4 are located near the C-terminus of the  $\beta$ -trunk domain and are fitted with the additional  $\beta$ -COP C-terminal helices. Densities 5 and 7 are located in the vicinity of the  $\delta$ -COP subunit and are fitted with the helices belonging to that subunit (see also figure 5). See also Figure 4 – Figure Supplement 1.

#### **Figure 5**

##### **$\delta$ -COP contains extra $\alpha$ -helices and contacts the Arf1 GTPase**

(A) A schematic representation of  $\delta$ -COP domain architecture. Predicted helices downstream of the  $\delta$ -COP longin domain are labeled “a”, “b”, “c”. (B) Arf1/ $\delta$ -COP/ $\beta$ -COP complex fitted into the EM density. The MHD  $\delta$ -COP is not shown here, since it is a part of the inter-triad linkages and is averaged out in the leaf structure. Residues 139-165 of helix b are modeled. Color scheme: EM density – transparent grey, Arf1 – pink,  $\beta$ -COP – dark green,  $\delta$ -COP – orange. The  $\delta$ -COP helix “a” is highlighted in yellow, and helix “b” in red. (C) The Arf1/ $\delta$ -COP/ $\beta$ -COP subcomplex fitted into the EM map illustrated unmasked at a lower isosurface level. (D) Arf1 in the GDP-bound state. The orientation of Arf1-GDP is equivalent to that in panel C. The position occupied by the  $\delta$ -COP helix b in the assembled coat is shown in semi-transparent red: it overlaps with helix N0. Color code: GDP – green, Arf1 – pink, Switch I – cyan, Switch II – purple, Interswitch – blue, helix N0 – green.

## Figure 6

### Site-directed photo-cross-linking

(A) Site-directed photo-cross-linking of chimeric  $\beta\delta$  subcomplexes and Arf1-GTP $\gamma$ S on Golgi-like liposomes, with photolabile aminoacids (Benzophenone (Bp)) in  $\delta$ -COP positions 156 or 159 (Ct $\beta$ 19-391-M $\delta$ T156Bp and Ct $\beta$ 19-391-M $\delta$ R159Bp). After irradiation with UV light, cross-linked products were analyzed by SDS-PAGE and western-blot with antibodies directed against Arf1 (left panel) and  $\delta$ -COP (right panel), and are marked with black asterisks. Bands of non-cross-linked proteins are marked with arrows (IB: immune blot). (B) Site-directed photo-cross-linking of Arf1-GTP $\gamma$ S with Bp in both positions 46 and 167 (Arf1-I46BpY167Bp) with coatomer on Golgi membranes. Cross-linked products were analyzed by SDS-PAGE and western blot with antibodies directed against  $\beta$ -COP,  $\delta$ -COP and Arf1. Black asterisks mark double cross-

1082 linked products linked by both photolabile residues, red asterisks mark single cross-  
 1083 linked products linked by Bp at either position 167 or 46 in Arf1. Non-cross-linked  
 1084 proteins are marked with arrows. **(C)** Ribbon model of the structure of a subcomplex of  
 1085 Ct $\beta$ 19-391 $\delta$ 1-159 with Arf. Residues involved in cross-linking are shown as spheres  
 1086 (orange-red for  $\delta$ -COP and purple for Arf1). In summary, photolabile amino acids  $\delta$ -  
 1087 COP156 and  $\delta$ -COP159 cross-linked to Arf1, photolabile Arf46 cross-linked to  $\beta$ -COP,  
 1088 and Arf167 cross-linked to  $\delta$ -COP. Mass-spectrometry cross-linking also identified a  
 1089 cross-link between Arf36 and  $\delta$ -COP142 (Dodonova et al., 2015). **(D)** Binding of  
 1090 N $\Delta$ 20CtArf1 GMPPNP to Ct $\beta\delta$  subcomplexes. Subcomplexes contained  $\beta$ 19-391-COP and  
 1091  $\delta$ -COP including helix a ( $\delta$ -COP1-137), or  $\delta$ -COP including helix b ( $\delta$ -COP1-175), or  $\delta$ -  
 1092 COP including helix c ( $\delta$ -COP1-243), or full-length  $\delta$ -COP, as indicated in the figure. Ct $\beta\delta$   
 1093 subcomplexes were immobilized on Strep-Tactin sepharose beads. Beads were  
 1094 incubated with purified N $\Delta$ 20CtArf complexed with GMPPNP or GDP. Pulldowns were  
 1095 analyzed by SDS-PAGE and western-blot. The gels were cut in two pieces. The lower  
 1096 piece was immuno-blotted with antibodies directed against Arf1 (lower panel). The  
 1097 upper part was used for coomassie staining to visualize COP subcomplexes (upper  
 1098 panel) (Note:  $\delta$ -COP fragments 1-137, 1-175 and 1-243 are not visible in the coomassie  
 1099 stained upper panel as they migrate into the part of the gel that was blotted for  
 1100 quantification of Arf1). **(E)** Quantification of the data depicted in D. As a control, binding  
 1101 of N $\Delta$ 20CtArf1 to  $\beta$ 19-391COP complexed with full length  $\delta$ -COP was analyzed in the  
 1102 presence of GDP (last column). Pulldowns were quantified using the Image-Studio  
 1103 software (Li-Cor Bioscience). Quantification was normalized to the  $\beta\delta$ -COP  
 1104 subcomplexes containing full-length  $\delta$ -COP with N $\Delta$ 20CtArf in its GMPPNP complexed



state. (means +/- SEM; n=3). See also Figure 6 – Figure Supplement 1 and Figure Supplement 2. Source data files for panel E are available in Figure 6-source data 1.

**Figure 6-source data 1.**

**Source data file for Figure 6E.**

This source data file contains the values obtained by quantifying the pulldown experiments (Figure 6D) using Image-Studio software, and presented in Figure 6E.

**Figure 7**

**Localization of ArfGAP2 within the COPI coat**

(A) COPI-ArfGAP2 leaf structure as grey isosurface. (B) COPI leaf structure (grey) and the difference map between the COPI and the COPI-ArfGAP2 structure (red). (C) The fit of COPI and the ArfGAP2 catalytic domain (PDB:2P57) into the COPI-ArfGAP2 leaf structure. The domain is located near the Arf1-nucleotide binding site (nucleotide is shown in green) (D) Surface representations of the COPI and ArfGAP2 molecular models, illustrating the position of the ArfGAP2 catalytic domain (red) in a niche within the assembled coat. See also Figure 7 – Figure Supplement 1.

## SUPPLEMENTARY FIGURES AND TABLES

### Figure 1 – Figure Supplement 1.

#### The structure of $\beta\delta$ -COP

(A) In the crystal structure,  $\beta$ - and  $\delta$ -COP share an extensive interface with a size of 1737 Å<sup>2</sup>. The longin domain,  $\beta$ -turn, and helix a of  $\delta$ -COP are part of the interface with the inside of the  $\beta$ -COP  $\alpha$ -solenoid, whereas helix b traverses the  $\alpha$ -solenoid along the connection of helices  $\alpha$ 7 and  $\alpha$ 6 of  $\beta$ -COP and protrudes away from the solenoid (see also Figure 1). F133 in helix a is in close contact with  $\alpha$ -solenoid residues Y114 and Y149. The side chain of  $\delta$ -M136 anchors the short region between helix a and b onto  $\beta$ . Additionally, the side-chain of  $\beta$ -N112 forms a hydrogen bond to the main-chain oxygen of  $\delta$ -M136. Other interactions are a salt bridge between  $\delta$ -E141, which is located at the beginning of helix b, and side chain of  $\beta$ -K71, and a hydrogen bond between side chains of  $\delta$ -S138 and  $\beta$ -E113. The side chains of I143, I146, and I147 in helix b are shown. These residues are essential for correct trafficking of HDEL motif containing cargo by COPI (Arakel et al., 2016). (B) In the crystal,  $\delta$ -COP helix b (red) is involved in a contact with helix b of a symmetry-related molecule (grey). The crystal contact comprises an interface area of 371 Å<sup>2</sup>. We cannot assess whether the crystal contact influences the orientation of helix b, but superposition of corresponding parts of  $\beta$ - and  $\delta$ -COP from the X-ray and cryo-ET structures (not shown) suggests that there is some flexibility in this region. (C) In the X-ray structure of  $\gamma\zeta$ -COP with bound Arf1 (Yu et al., 2012), helices  $\alpha$ 2,  $\alpha$ 4, and  $\alpha$ 6 of  $\gamma$ -COP interact with Arf1.  $\gamma\zeta$ -COP helix  $\alpha$ 4 shows standard  $\alpha$ -helical geometry from residues T63 to T74 and ends in a  $3_{10}$  helical conformation (K75-Q78). In contrast, in our structure of  $\beta\delta$ -COP, downstream of residue I61,  $\beta$ -COP helix  $\alpha$ 4 is shorter, and between residues I61 and F64 it forms a widened helix with a rare  $\pi$ -

helix-like geometry. To highlight the quality of the structure in this region, electron density at contour level  $4\sigma$  from the anomalous diffraction experiment with SeMet-labelled protein is shown for the two methionine residues M59 and M66.

(D) The main chain oxygen of  $\beta$ -COP F64 in helix  $\alpha 4$  forms a hydrogen bond with the amino group of the side chain of K37, which is located in the preceding helix  $\alpha 3$ . 2mFo-DFc electron density is shown at contour level  $1\sigma$ . (E) The x-ray structure of *Chaetomium thermophilum*  $\beta\delta$ -COP can be fit as a rigid body into the cryo-ET map of the *Mus musculus* COPI leaf at 9.2 Å resolution, indicating a high degree of structural conservation.  $\beta$ -COP is shown in dark green,  $\delta$ -COP in orange.

## **Figure 2 – Figure Supplement 1.**

### **The COPI coat structure at 9 Å or 13 Å resolution**

(A) FSC curves (gold standard) for the leaf structure before (red) and after local alignments (blue). For the locally aligned structure the resolution at the 0.143 criterion is 9.2 Å, at 0.5 FSC – 11.1 Å. (B) The local resolution within the leaf structure ranges from 12.7 Å (red) to 8.7 Å (blue.) After local alignments, the resolution ranges from 11.5 to 8.1 Å. The color range for the locally aligned leaf is the same as for the initial structure for comparison purposes. (C) EM reconstruction of the COPI triad at 13 Å resolution (Dodonova et al., 2015). The triad is colored based on the radial distance from the membrane (from red to blue). (D) The locally aligned EM reconstruction of the leaf at 9.2 Å resolution for comparison to the 13 Å structure. At 9.2 Å resolution,  $\alpha$ -helical densities are clearly resolved. The leaf structure is colored based on the radial distance from the membrane. The orange dashed line indicated the cutting plane of the orthoslice shown in the panel B. (E) Slices through the EM density of the 13 Å resolution (left panel), and the 9.4 Å resolution coat structure before the local alignment (right

panel) for comparison. The cutting plane is shown with a dashed line in the panel C and D. Note the clear outlines of the  $\alpha$ -helical densities in the structure in the right panel.

## **Figure 2 – Figure Supplement 2.**

### **COPI coat linkages**

(A) The structures of the COPI linkages I, II, III, IV (see (Dodonova et al., 2015) for description). The upper panels show the linkage structures colored based on the radial distance from the membrane (from red to blue). The lower panels show the structural models generated by fitting protein structures into the linkage EM maps. The subunits that mediate the main contacts in the linkages are  $\alpha\epsilon$ -COP (blue and cyan), and  $\delta$ -COP MHD (orange). (B) The FSC curves for the linkage structures. Resolution at the 0.143 FSC is 15-17 Å for the linkages I, II, IV and 30 Å for linkage III, which is the least abundant of all linkages. (C) The inter-leaf contacts within linkages I, II and IV. Within linkage IV there are two non-symmetry related leaves: in one  $\beta$ Arf1 abuts a neighboring triad (labeled linkage IV), in the other  $\beta$ Arf1 protrudes towards the center of the linkage (labeled linkage IV\*). The contacts made by Arf1 are marked with pink asterisks, the contact between  $\alpha$ -COP and  $\gamma$ -COP is marked with a blue asterisk. The boundaries between different leafs are shown with orange lines, the boundaries between triads are shown with a dotted black line. Note that the molecular contacts made by  $\beta$ Arf1 in the linkages I, II and IV are almost identical, while  $\beta$ Arf1 in Linkage IV\* makes fewer contacts. Isosurface threshold is  $1\sigma$ .

## **Figure 2 – Figure Supplement 3.**

**The spatial relationships between components of the coat are illustrated here as an aid to the reader.**

The vesicle used for illustration is from our previously published work (Dodonova et al., 2015). **(A)** A composite representation of a complete COPI vesicle. The membrane is shown in grey. The color scheme of the subunits is shown in the panel. **(B)** A composite representation of a COPI vesicle, showing only the trunk adaptor domains of  $\gamma$ -COP (light green) and  $\beta$ -COP (dark green). **(C)** A composite representation of a COPI vesicle, showing only the core outer-coat subunits  $\alpha$ -COP (dark blue) and  $\beta'$ -COP (light blue). **(D)** A composite representation of a COPI vesicle, showing the trunk domains of  $\gamma$ -COP (light green) and  $\beta$ -COP (dark green) and the associated Arf1 molecules –  $\gamma$ Arf (pink),  $\beta$ Arf (purple). **(E)** A composite representation of a COPI vesicle, showing only the Arf1 molecules –  $\gamma$ Arf (pink),  $\beta$ Arf (purple).

#### **Figure 4 – Figure Supplement 1.**

##### **Identification of extra densities in the COPI EM map**

**(A)** The model of the COPI leaf was flexibly fitted into the EM map, and transformed into an EM density with 9 Å resolution (chimera molmap). The model density was subtracted from the initial EM map and the resulting difference map is shown in green and yellow. **(B)** The volume distribution of the extra densities (in voxels). The seven largest densities are highlighted in green in the plot and in the panel A. Source data files for panel B are available in Figure 4-S1-source data 1.

#### **Figure 4-S1-source data 1.**

##### **Source data file Figure 4 Figure Supplement 1B.**

This source file contains the volume values (in voxels) for the unoccupied “extra” densities shown in Figure 4 - Figure Supplement 1B, as determined using Chimera.

**Figure 6 – Figure Supplement 1.**

**Mass-spectrometry derived cross-links from newly positioned COPI domains**

Cross-linked peptides from the newly assigned regions were mapped within the COPI leaf model (cyan lines).  $\delta$ -COP helix b (red) was cross-linked to  $\beta$ -COP (dark green), to Arf1 (pink), and to  $\beta'$ -COP (light blue). The  $\delta$ -COP helix C (orange) was cross-linked to  $\delta$ -COP MHD (orange). Note that  $\delta$ -COP MHD is located in the linkage, and not in the COPI leaf, so it is not enclosed by the EM density.

**Figure 6 – Figure Supplement 2.**

**Site-directed photo-cross-links**

Site-directed photo-cross-linking of Arf1-I46BpY167Bp (loaded with GTP $\gamma$ S) to coatomer on (A) Golgi membranes, or (B) Golgi-like liposomes. Cross-links were analyzed by SDS-PAGE and western blot with antibodies directed against all coatomer subunits and against Arf1. Black asterisks mark double cross-link products obtained with photolabile Bp at positions 46 and 167 in Arf1, red asterisks mark single cross-link products with photolabile Bp at positions 46 or 167 in Arf1 (Sun et al., 2007). Bands representing the non-cross-linked proteins are marked with arrows.

**Figure 7 – Figure Supplement 1.**

**Position of the ArfGAP2 catalytic domain within in the COPI coat**

(A) The COPI-ArfGAP2 dataset was split into two distinct classes by multireference-based alignment and classification. Slice through the EM density of the class 1 from the COPI-ArfGAP2 dataset, which contains the extra density corresponding to the ArfGAP2 catalytic domain. The slicing plane is orthogonal to the membrane. The additional density is marked with an arrow. (B) Slice through the EM density of the class 2 from

the COPI-ArfGAP2 dataset, which did not contain any additional densities in comparison with the control structure. The slicing plane is orthogonal to the membrane. **(C)** Position of the catalytic domain (red) and ankyrin domain (cyan) based on the Arf6-ASAP3 structure (PDB:3LVQ). The Arf1 from the Arf6-ASAP3 structure was superimposed with the Arf1 in our model. Note, that the ankyrin domain clashes with the  $\beta$ -COP (dark green). The existence of a small COPI-niche suggests that there may be sterical selection for the type of ArfGAP protein interacting with Arf1 in the context of the COPI coat. **(D)** Position of the catalytic domain (red) based on the Arf1-ArfGAP1 structure, where Arf1 in the Arf1-ArfGAP structure was superimposed on the Arf1 in our model (Goldberg, 1998). Note, that the catalytic domain is far from the nucleotide site (nucleotide is shown in green). Color scheme: Arf1 - pink,  $\gamma$ -COP – light green,  $\beta$ -COP - dark green,  $\zeta$ -COP - yellow,  $\delta$ -COP - orange,  $\beta'$ -COP - light blue,  $\alpha$ -COP – dark blue. Compare panels C and D with our structural model illustrated in Figure 7C).

## **Table 1.**

### **X-ray data collection and refinement statistics**

Each dataset was collected from one single crystal. Statistics for the highest-resolution shell are shown in parentheses.

## **Table 2.**

### **Cross-links from newly assigned COPI domains**

The mass-spectrometry cross linking data is part of a previously published dataset (Dodonova et al., 2015). The distances between lysine pairs for which cross-links were observed were measured for our structural model. If the measured distance was below 35 Å, it satisfied the distance criteria.

1273

1274   **Video 1.**

1275   **The COPI coat structure and model.**

1276   A tour through the COPI structure highlighting key features of the coat discussed in this

1277   study.

1278



1279 Table 1. Data collection and refinement statistics.  
1280

Dataset	SeMet	native
Wavelength [Å]	0.97264	0.97625
Resolution range [Å]	46.40 - 2.70 (2.83 - 2.70)	32.05 - 2.57 (2.66 - 2.57)
Space group	C 2 2 2 <sub>1</sub>	C 2 2 2 <sub>1</sub>
Unit cell [Å]	a=138.22;b=176.40;c=62.60 (90°,90°,90°)	a=137.59;b=177.47;c=62.72 (90°,90°,90°)
Total reflections	185546 (6684)	165162 (19393)
Unique reflections	17882 (1081)	24909 (2440)
Multiplicity	10.4 (6.2)	6.6 (6.5)
Completeness [%]	83.2 (39.3)	99.9 (99.9)
Mean I/sigma(I)	20.8 (1.9)	15.1 (1.3)
R <sub>merge</sub>	0.069 (0.819)	0.068 (1.391)
R <sub>pim</sub>	0.031 (0.523)	0.031 (0.644)
CC1/2	0.999 (0.620)	0.999 (0.502)
# Reflections used in refinement		24904 (2439)
# Reflections used for R <sub>free</sub>		1228 (119)
R <sub>work</sub>		0.1748 (0.2802)
R <sub>free</sub>		0.2273 (0.3185)
Number of non-hydrogen atoms		4213
in Macromolecules		4169
Protein residues		521
RMSD on bonds [Å]		0.009
RMSD on angles [°]		1.19
Ramachandran favored [%]		97
Ramachandran allowed [%]		2.5
Ramachandran outliers [%]		0
Average B-factor [Å <sup>2</sup> ]		88.27
Macromolecules [Å <sup>2</sup> ]		88.42
Solvent [Å <sup>2</sup> ]		74.47

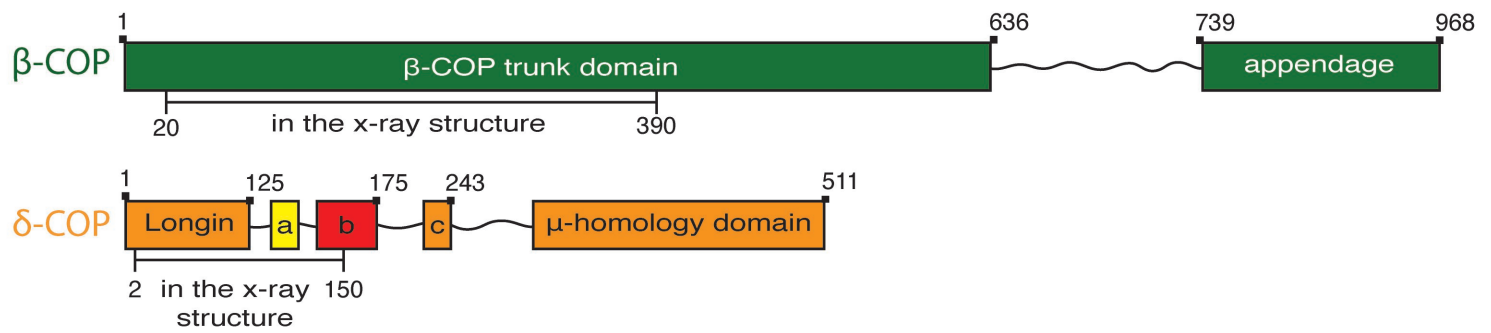
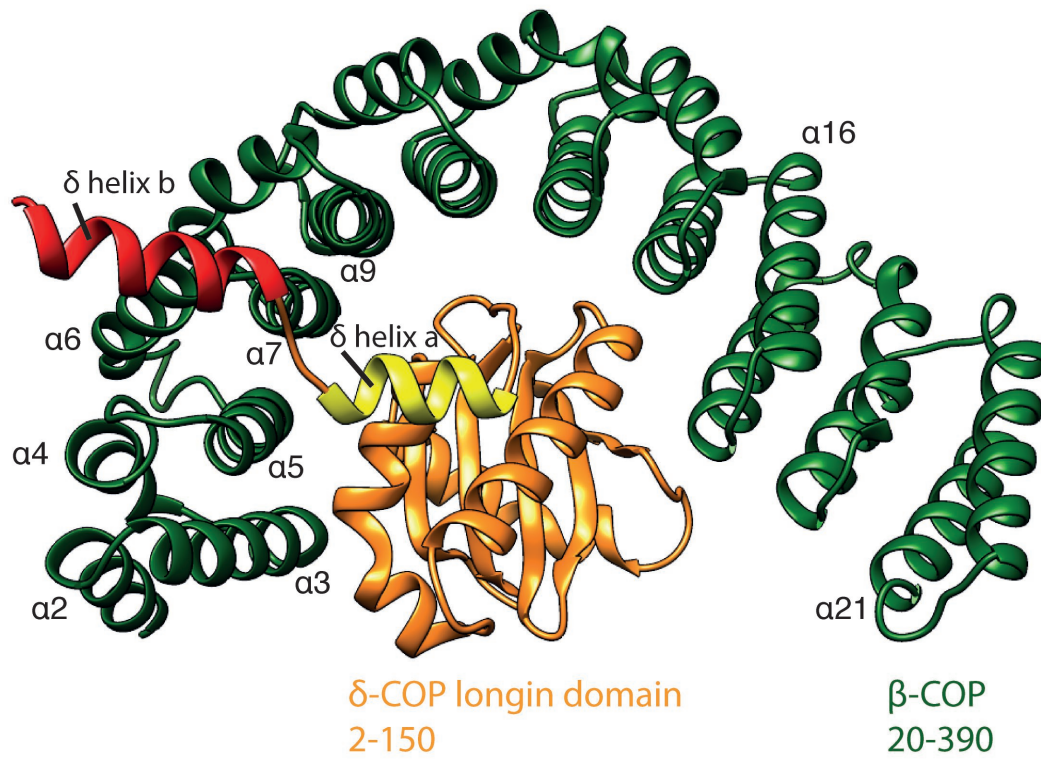
1281

1282

1283 Table 2. Cross-links from newly assigned COPI domains  
1284

#	Subunit 1	a.a.	Domain 1	Subunit 2	a.a.	Domain 2	ID score	Distance, nm
1	$\beta'$	856	C-terminal part	$\alpha$	1180	CTD	21,25	–
2	$\beta'$	871	C-terminal part	$\alpha$	1180	CTD	24,46	–
3	$\beta$	617	loop near extra helix	$\zeta$	39	core	27,11	–
4	$\beta$	618	loop near extra helix	$\zeta$	39	core	24,31	–
5	$\beta$	666	linker	$\beta$	891	appendage	46,62	–
6	$\beta$	671	linker	$\beta'$	779	$\alpha$ -solenoid	32,34	–
7	$\beta$	671	linker	$\beta$	891	appendage	47,02	–
8	$\beta$	694	linker	$\alpha$	750	$\alpha$ -solenoid	30,68	–
9	$\beta$	718	linker	$\beta'$	779	$\alpha$ -solenoid	28,42	–
10	$\beta$	718	linker	$\alpha$	671	$\alpha$ -solenoid	38,19	–
11	$\beta$	718	linker	$\alpha$	698	$\alpha$ -solenoid	36,71	–
12	$\beta$	718	linker	$\beta$	324	trunk	38,14	–
13	$\beta$	718	linker	$\beta$	801	appendage	35,16	–
14	$\beta$	718	linker	$\beta$	891	appendage	35,1	–
15	$\beta$	727	linker	$\alpha$	676	$\alpha$ -solenoid	25,12	–
16	$\delta$	142	extra helix b	Arf1	36	core	28,0	22
17	$\delta$	142	extra helix b	$\beta$	122	trunk	36,07	16
18	$\delta$	142	extra helix b	$\beta'$	69	1 <sup>st</sup> $\beta$ - propeller	31,54	60
19	$\delta$	156	extra helix b	$\beta$	122	trunk	31,37	24
20	$\delta$	156	extra helix b	$\beta$	47	trunk	26,7	27
21	$\delta$	164	near extra helix b	$\beta$	122	trunk	30,62	–
22	$\delta$	164	near extra helix b	Arf1	36	core	33,33	–
23	$\delta$	224	linker	$\beta'$	615	$\alpha$ -solenoid	18,14	–
24	$\delta$	227	linker	$\delta$	309	MHD	33,02	–
25	$\delta$	231	linker	$\delta$	309	MHD	33,28	–
26	$\delta$	233	extra helix c	$\beta'$	627	$\alpha$ -solenoid	29,71	26
27	$\delta$	233	extra helix c	$\delta$	243	extra helix c	38,83	16
28	$\delta$	233	extra helix c	$\delta$	256	linker	35,08	–
29	$\delta$	233	extra helix c	$\delta$	263	MHD	39,82	36
30	$\delta$	241	extra helix c	$\beta'$	627	$\alpha$ -solenoid	24,78	18
31	$\delta$	243	extra helix c	$\delta$	263	MHD	24,59	23
32	$\delta$	256	linker	$\delta$	263	MHD	31,39	–
33	$\delta$	256	linker	$\delta$	309	MHD	33,11	–
34	$\delta$	256	linker	$\delta$	322	MHD	24,85	–

1285



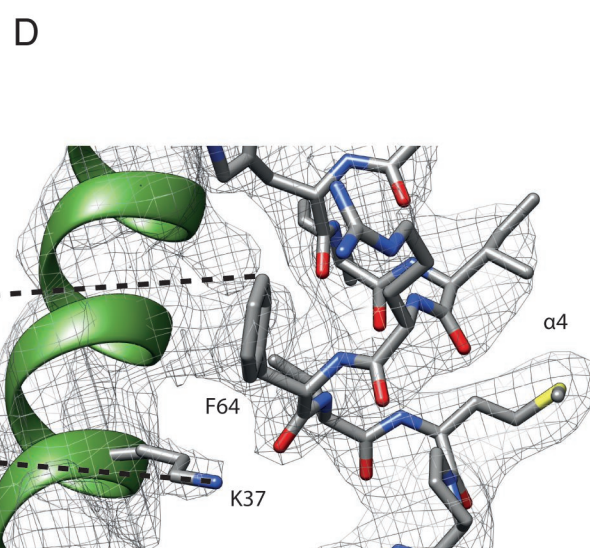
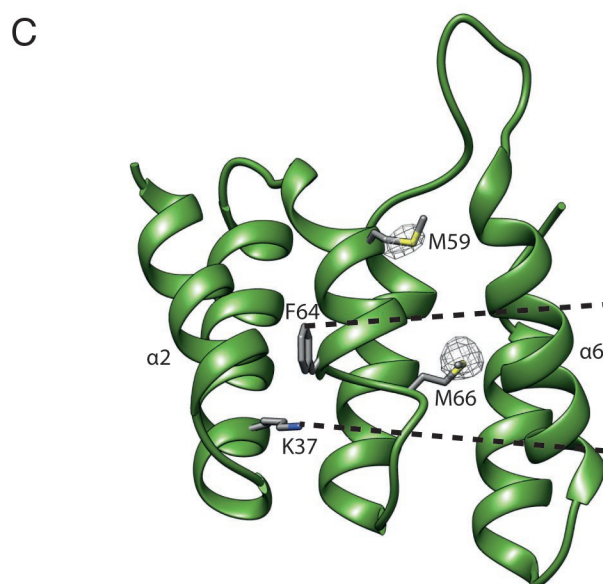
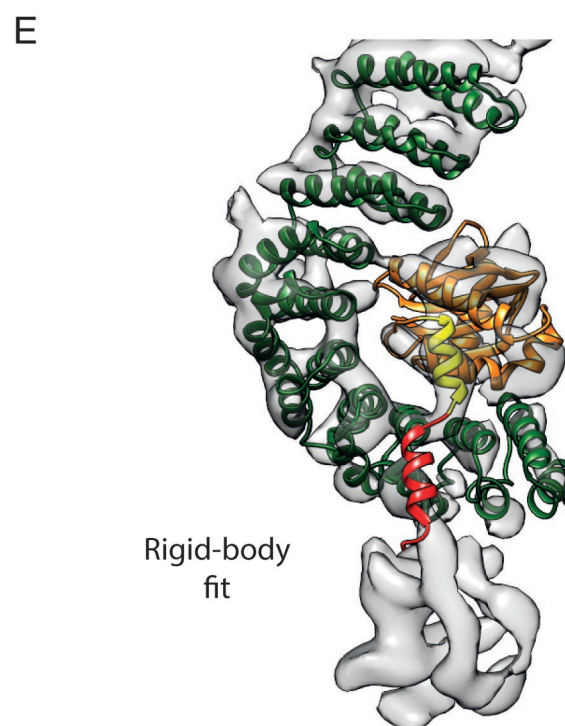
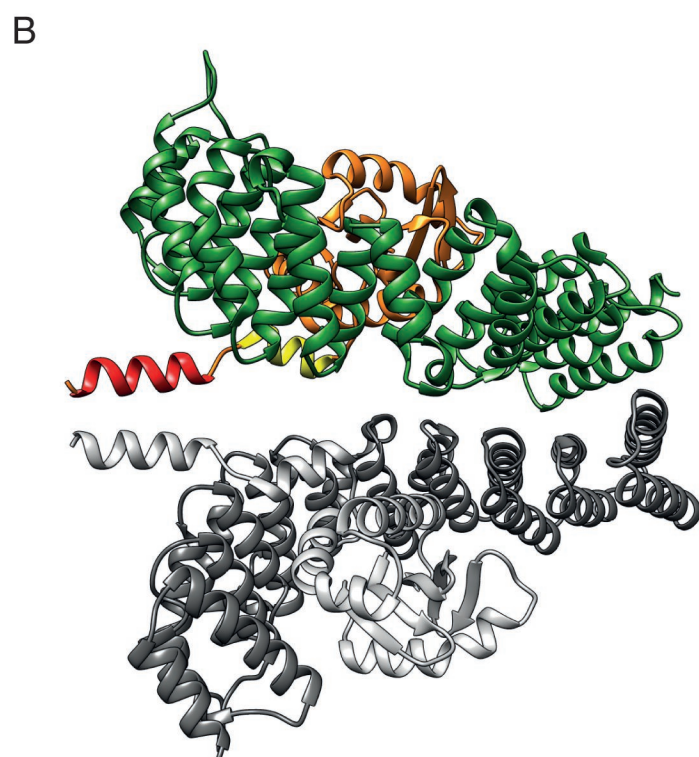
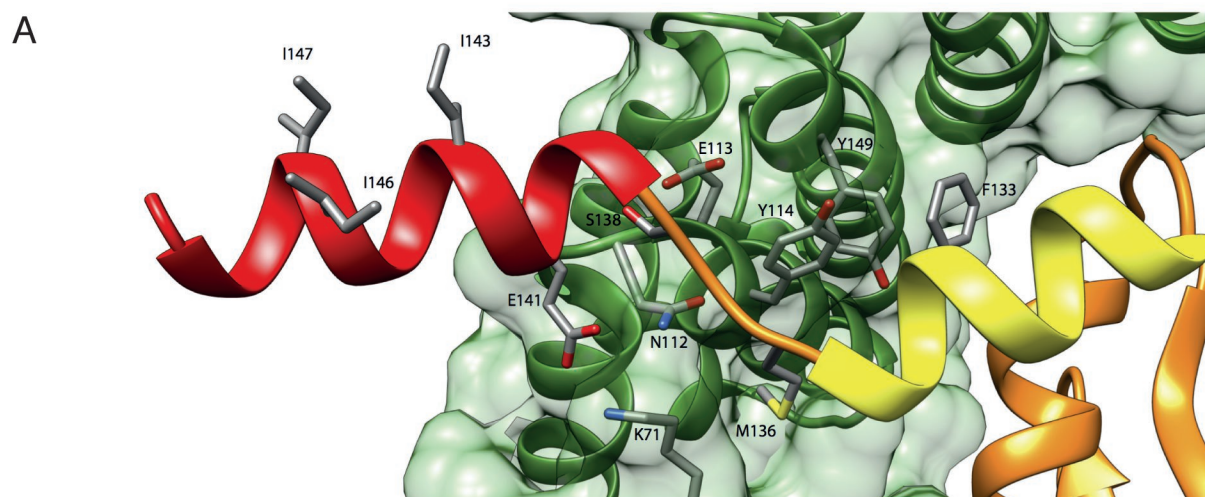
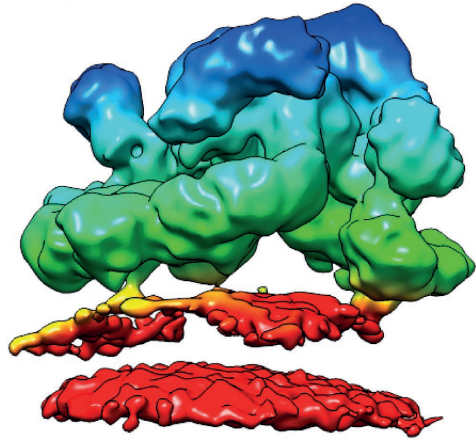
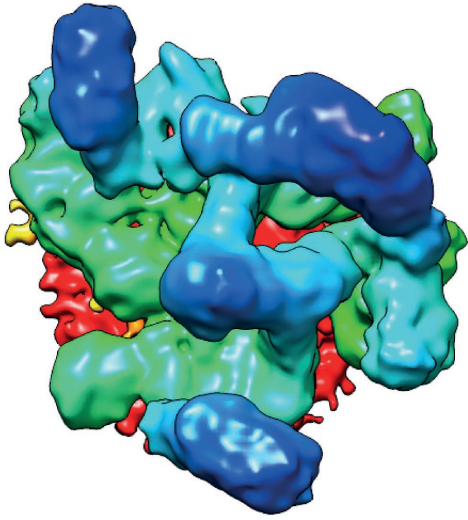


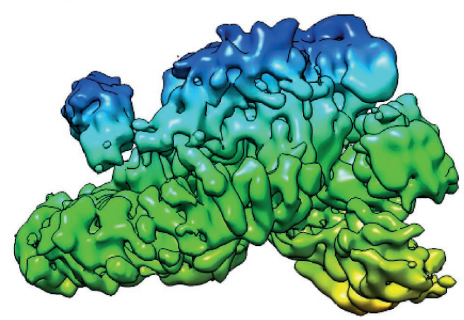
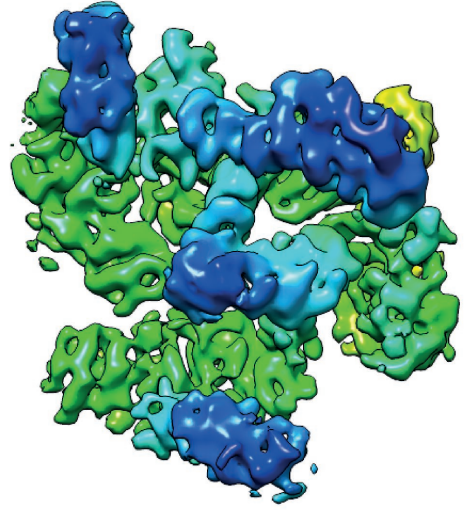


Figure 2

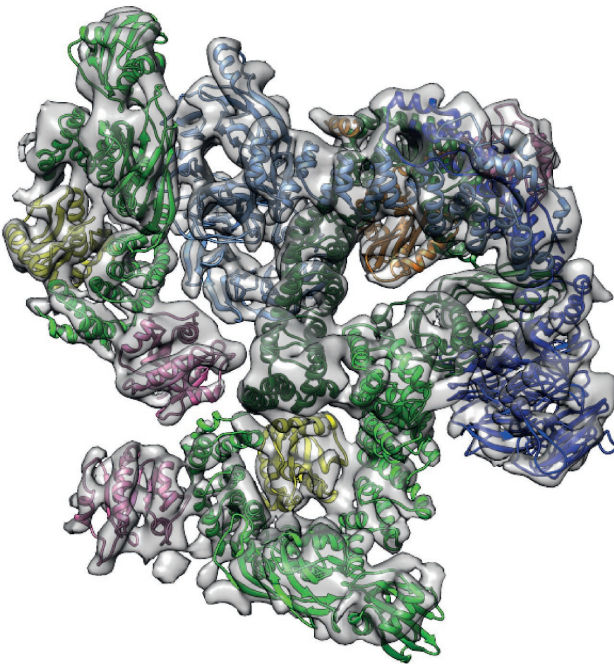
A



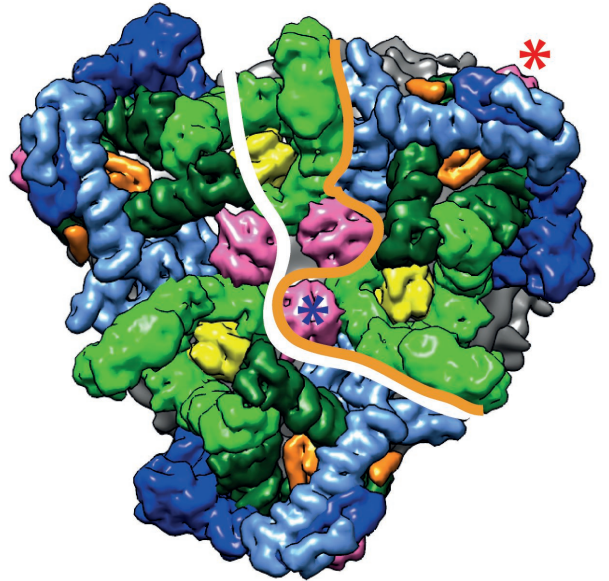
B



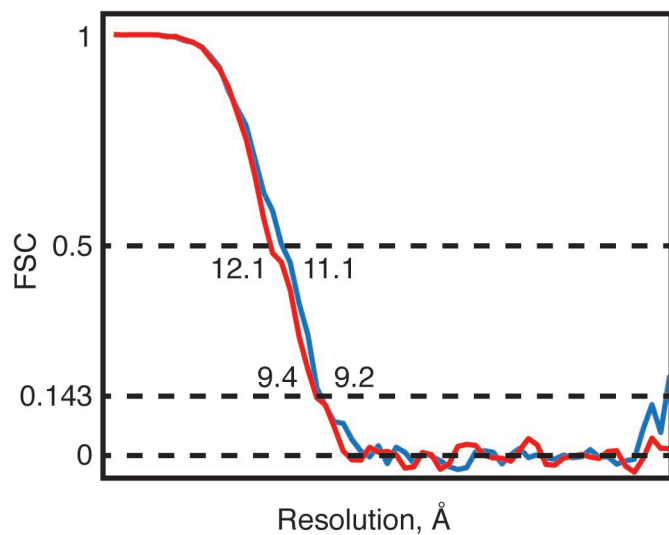
C



D

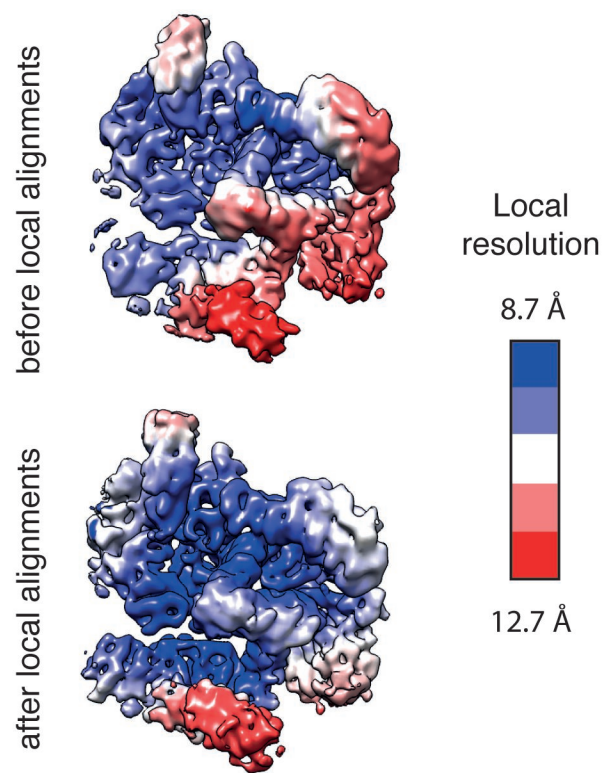


A

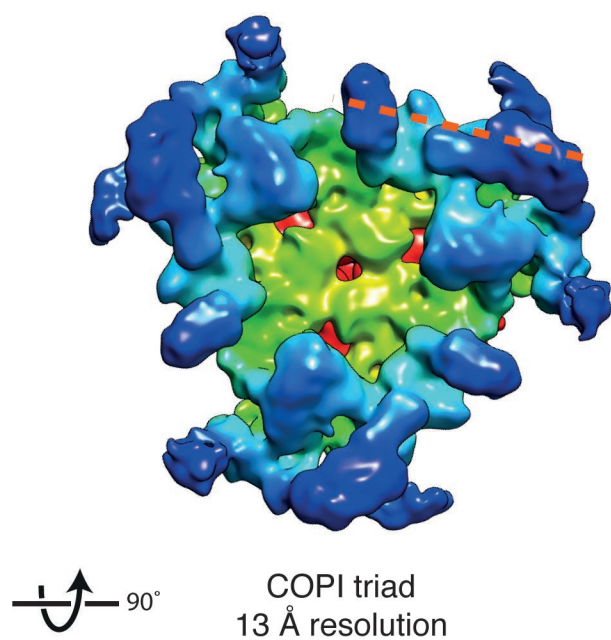


B

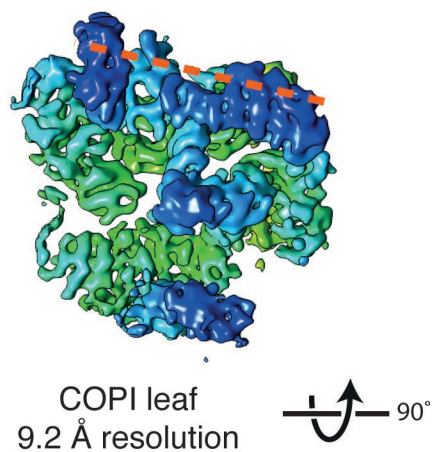
Figure 2 – Figure Supplement 1



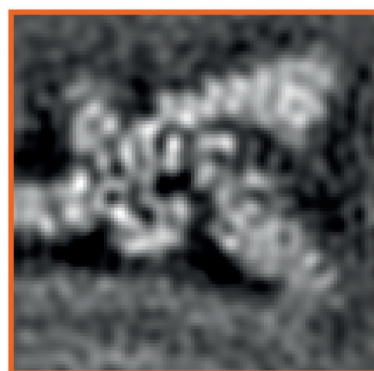
C



D

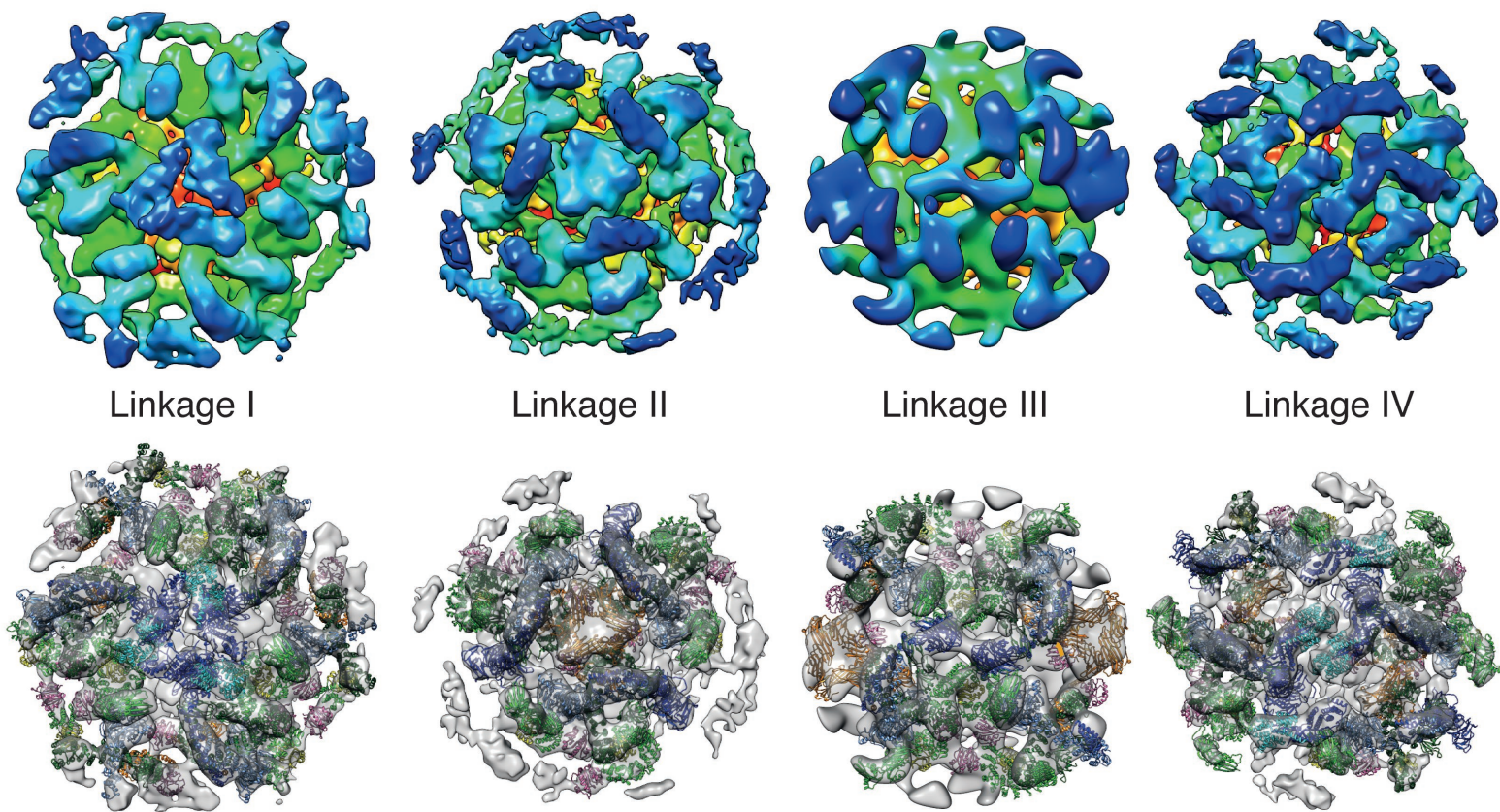


E

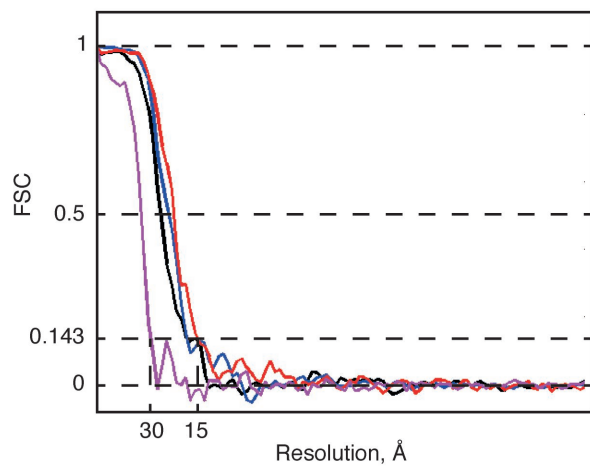




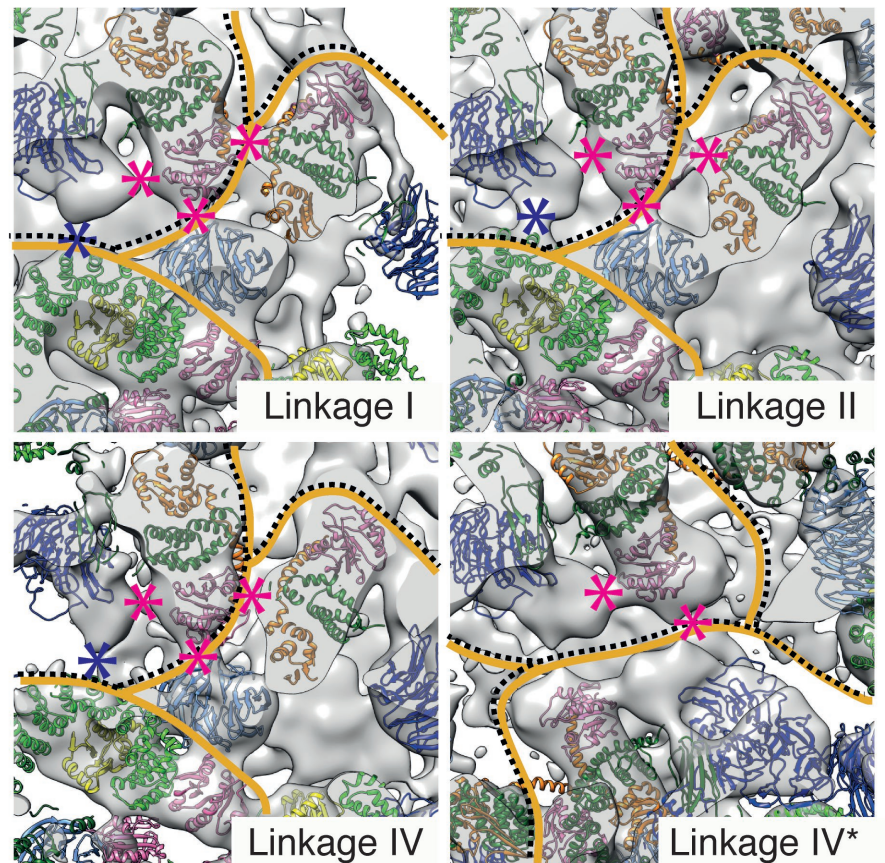
A



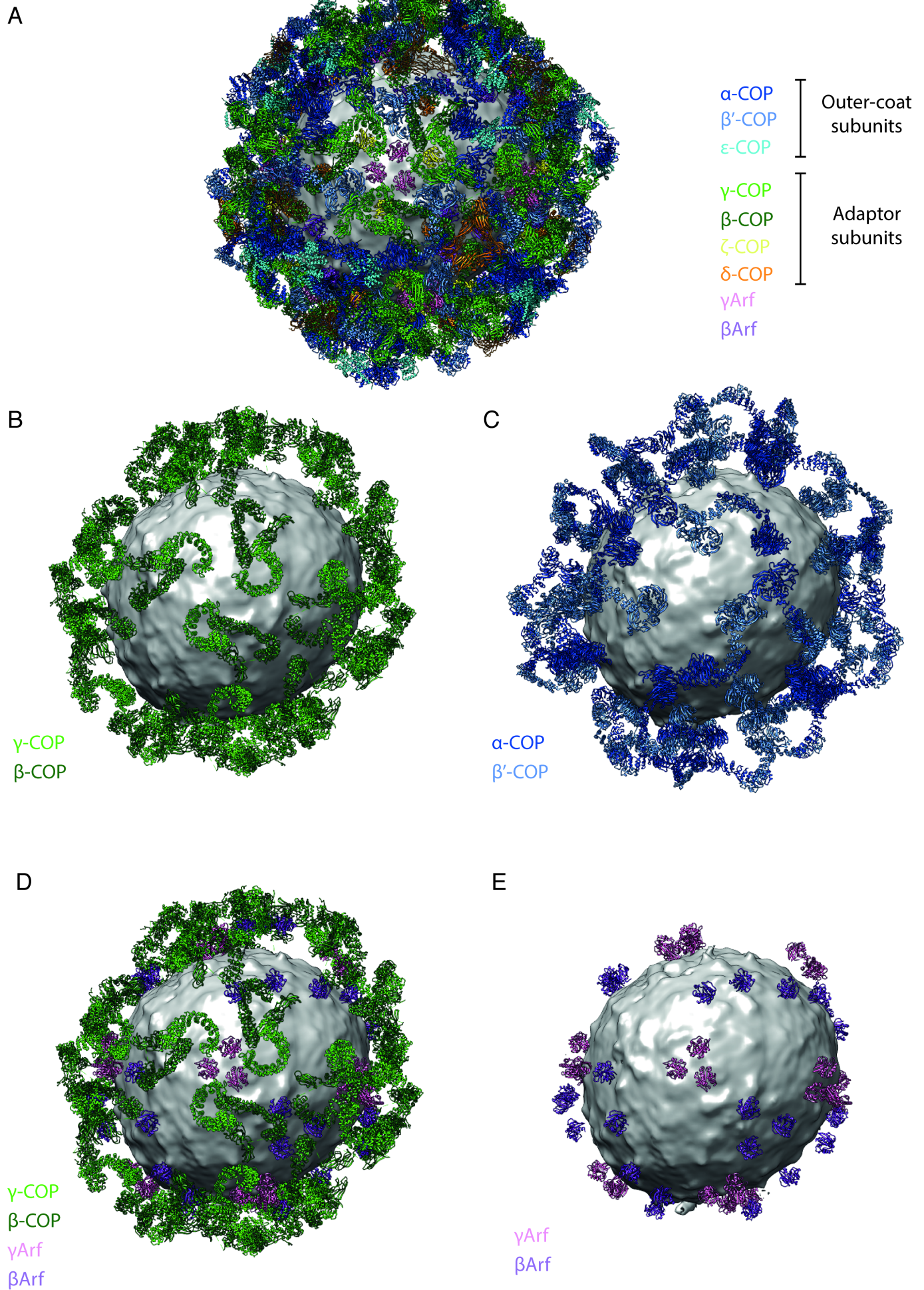
B



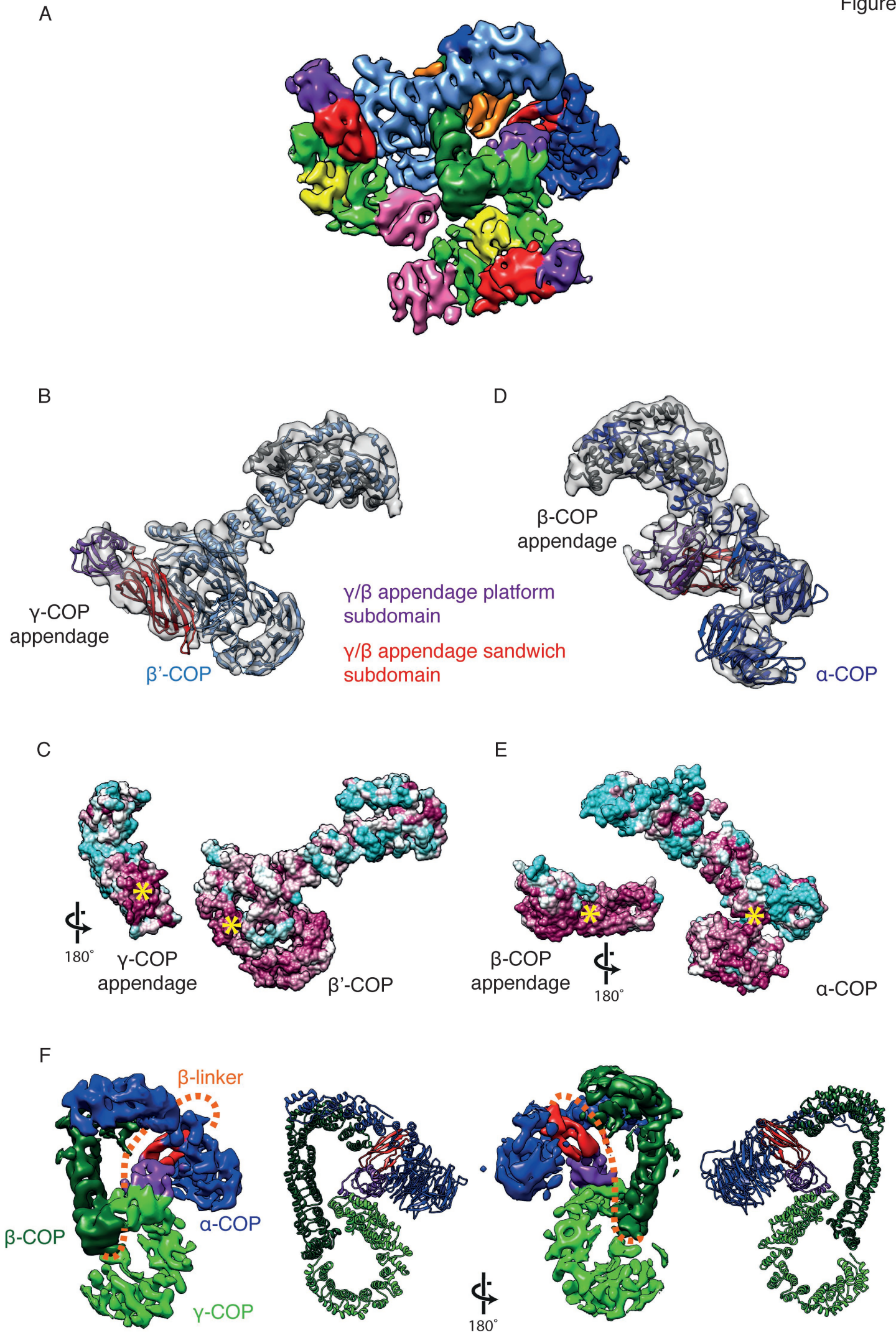
C





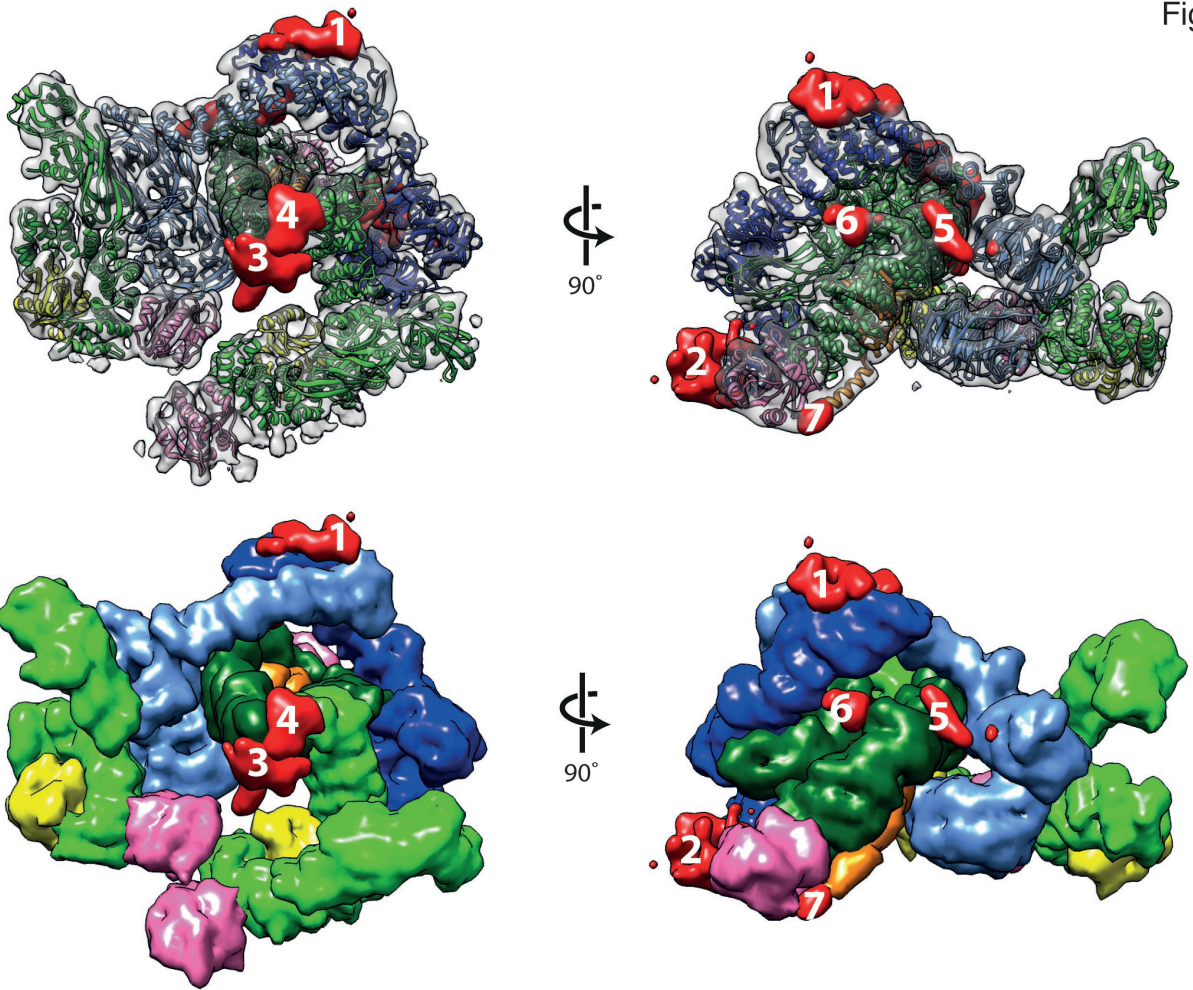




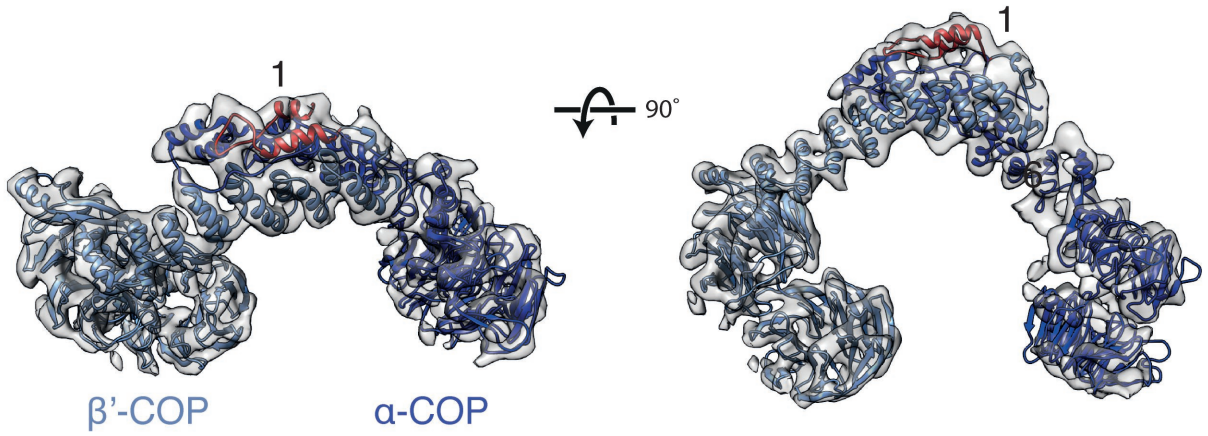




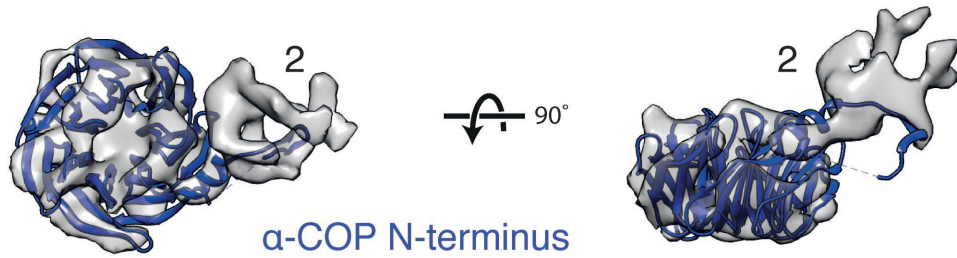
A



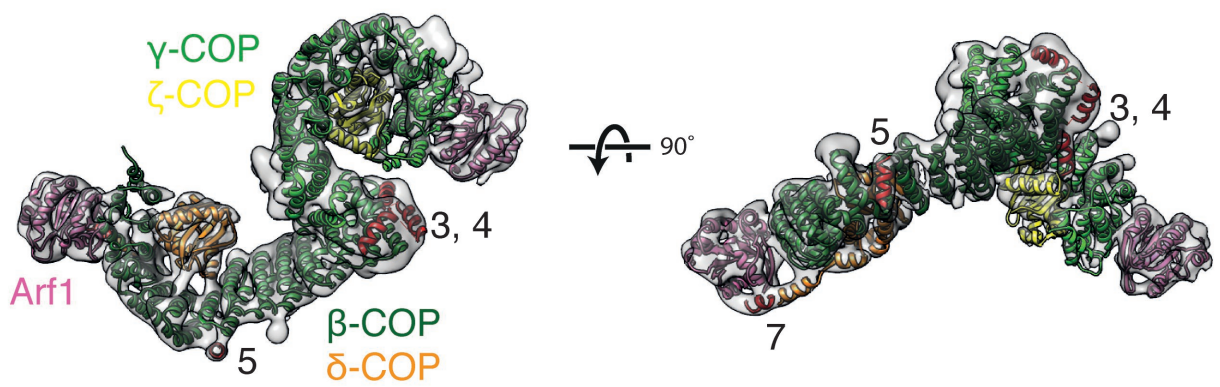
B



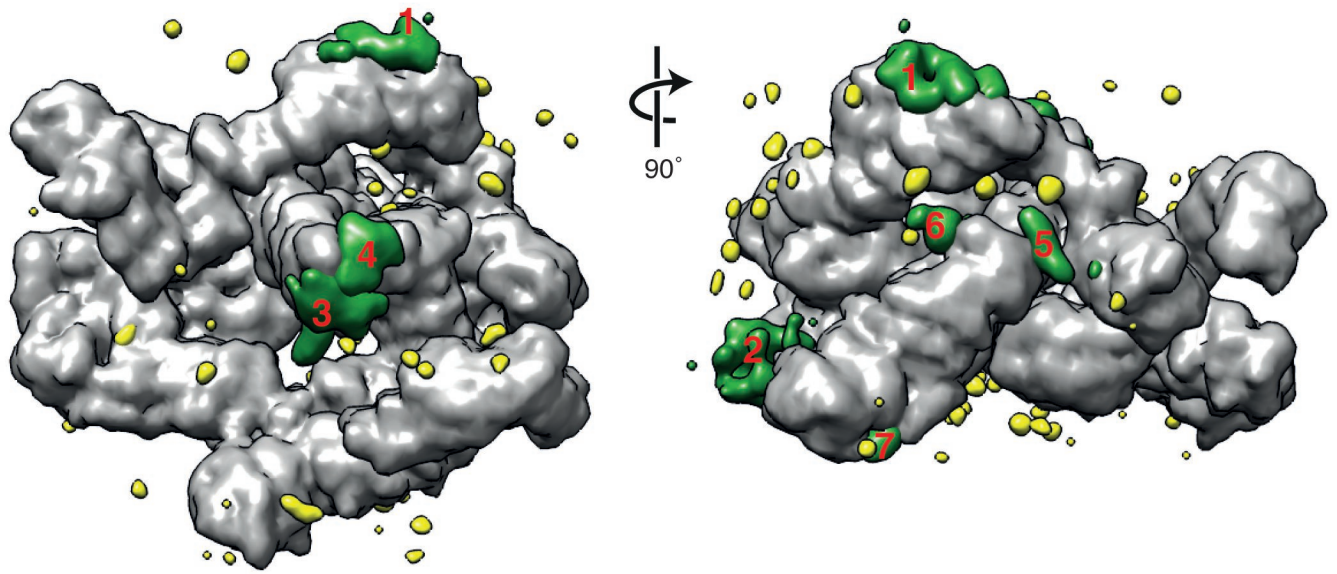
C



D



A



B

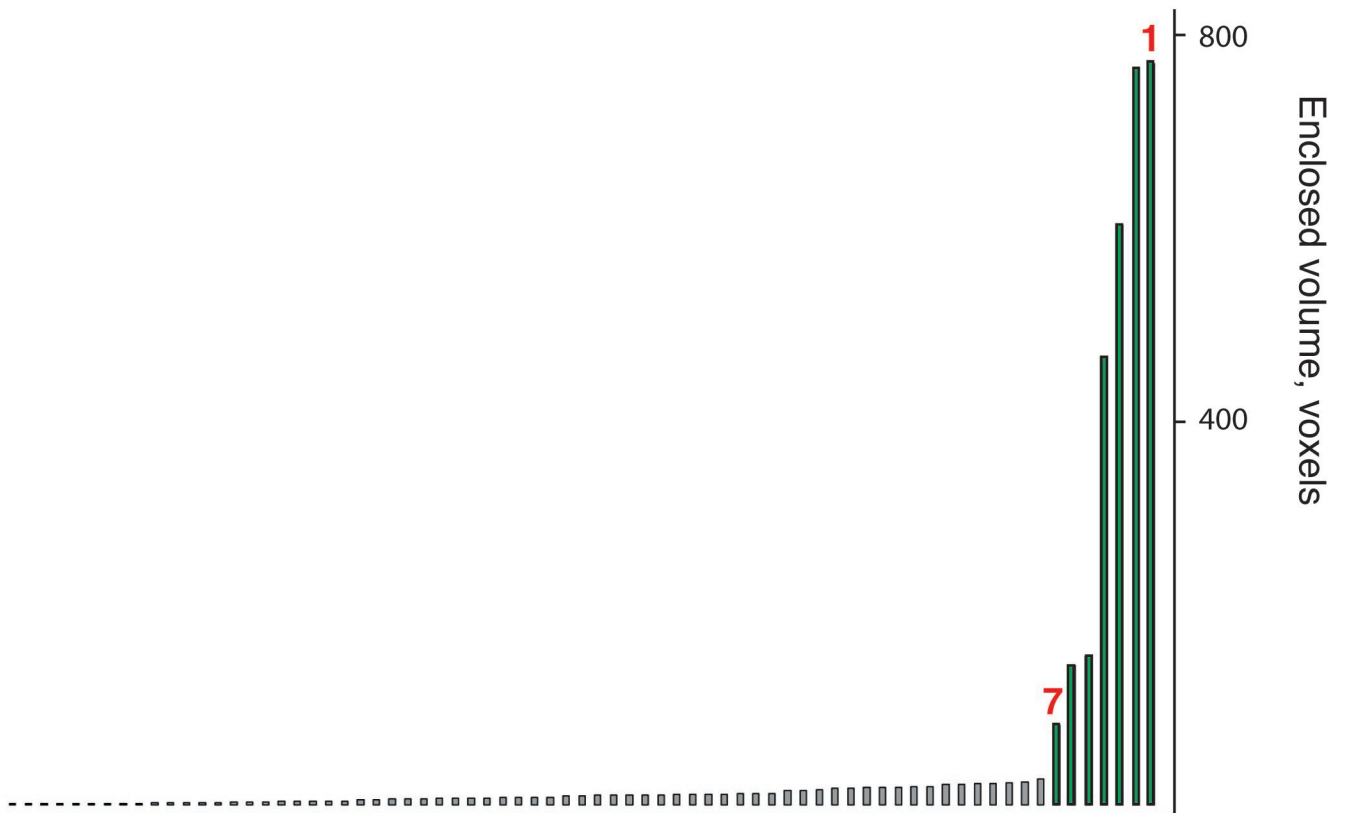




Figure 5

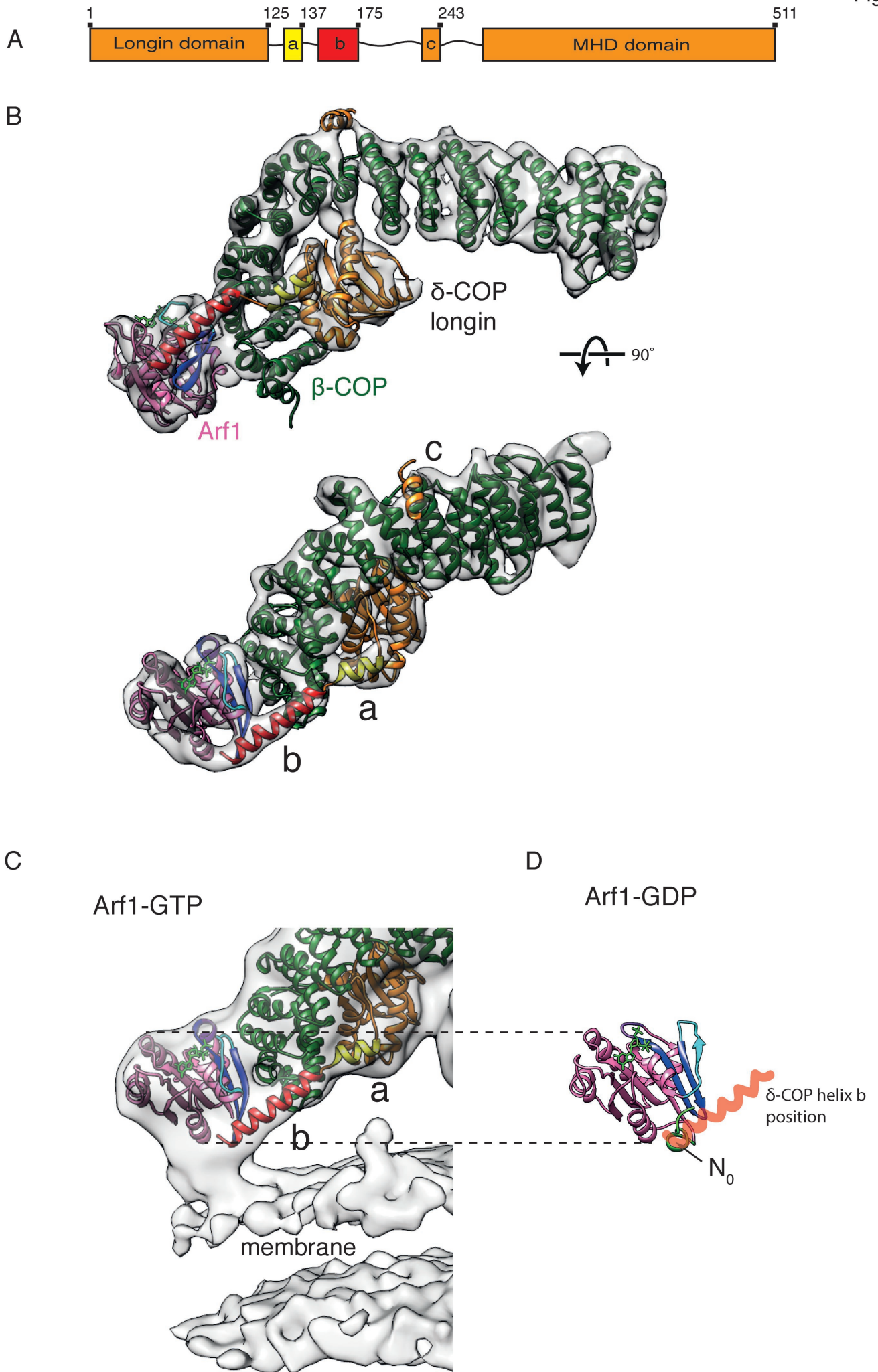
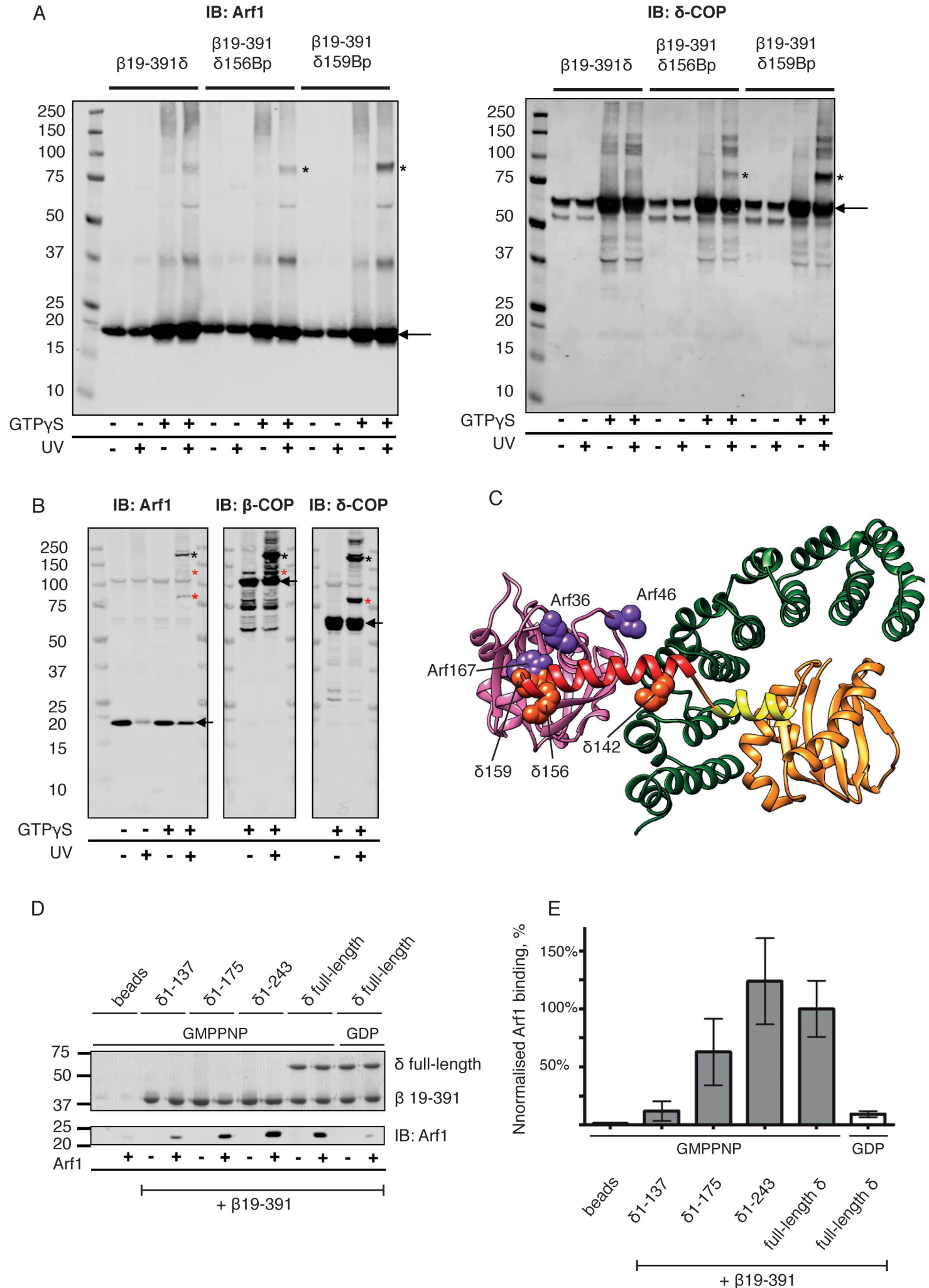
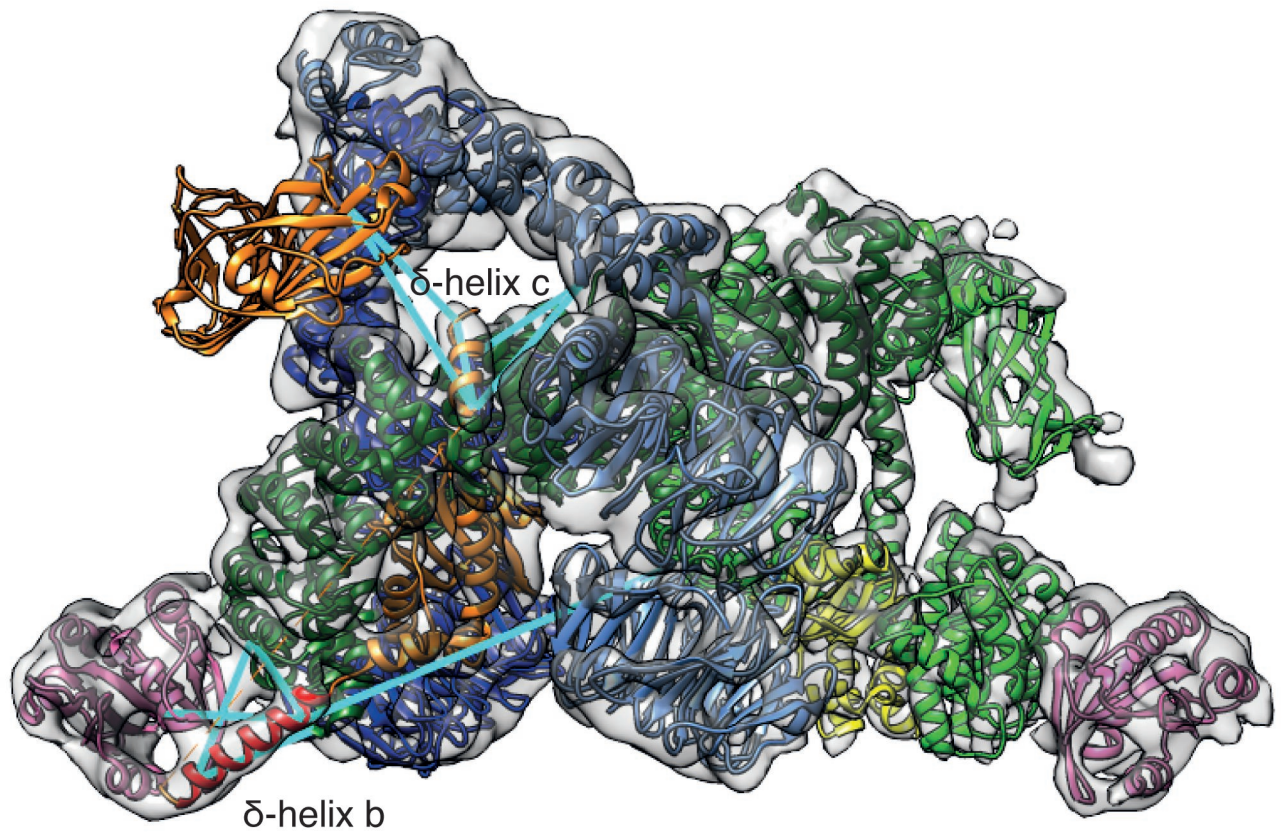
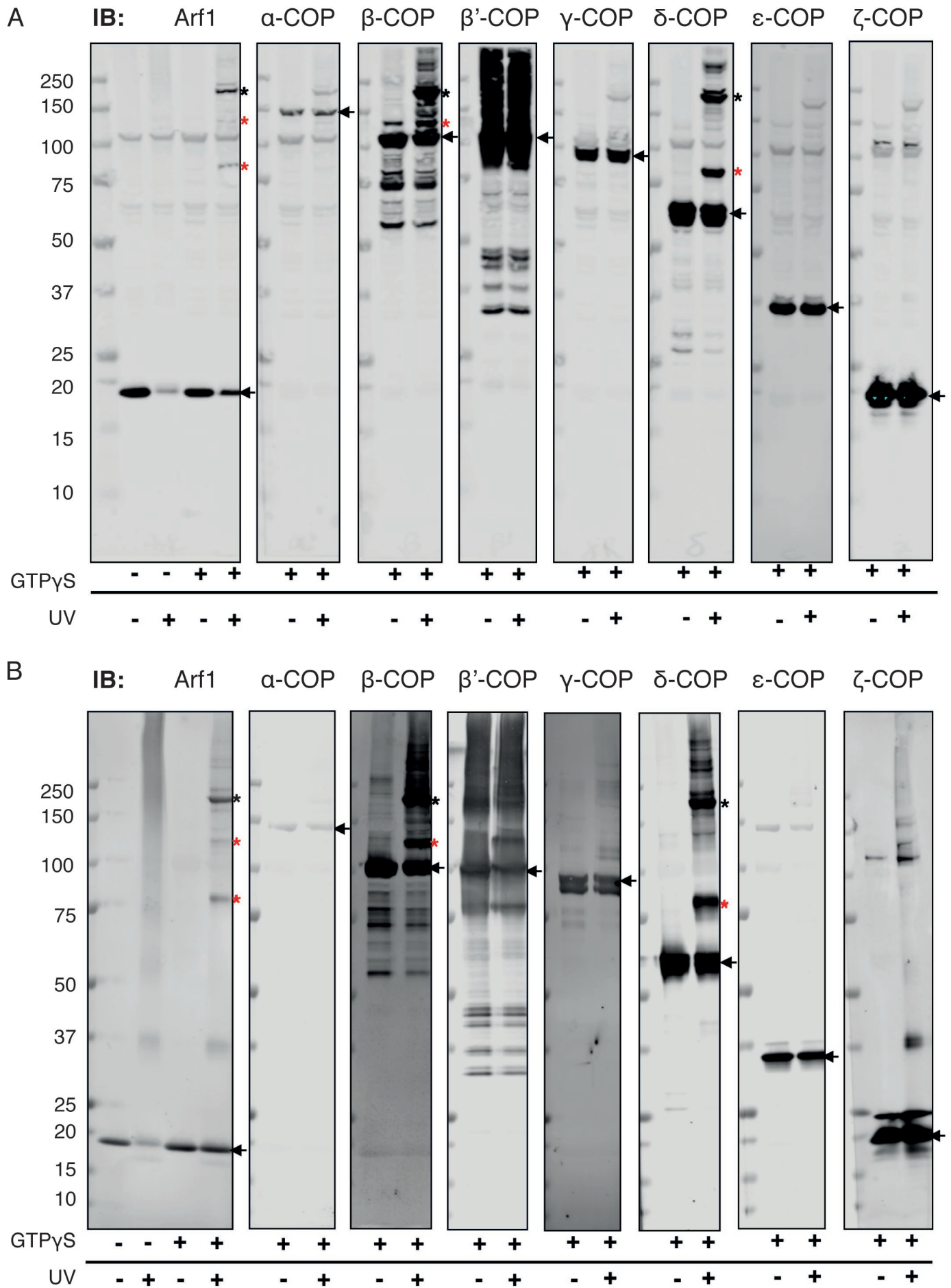


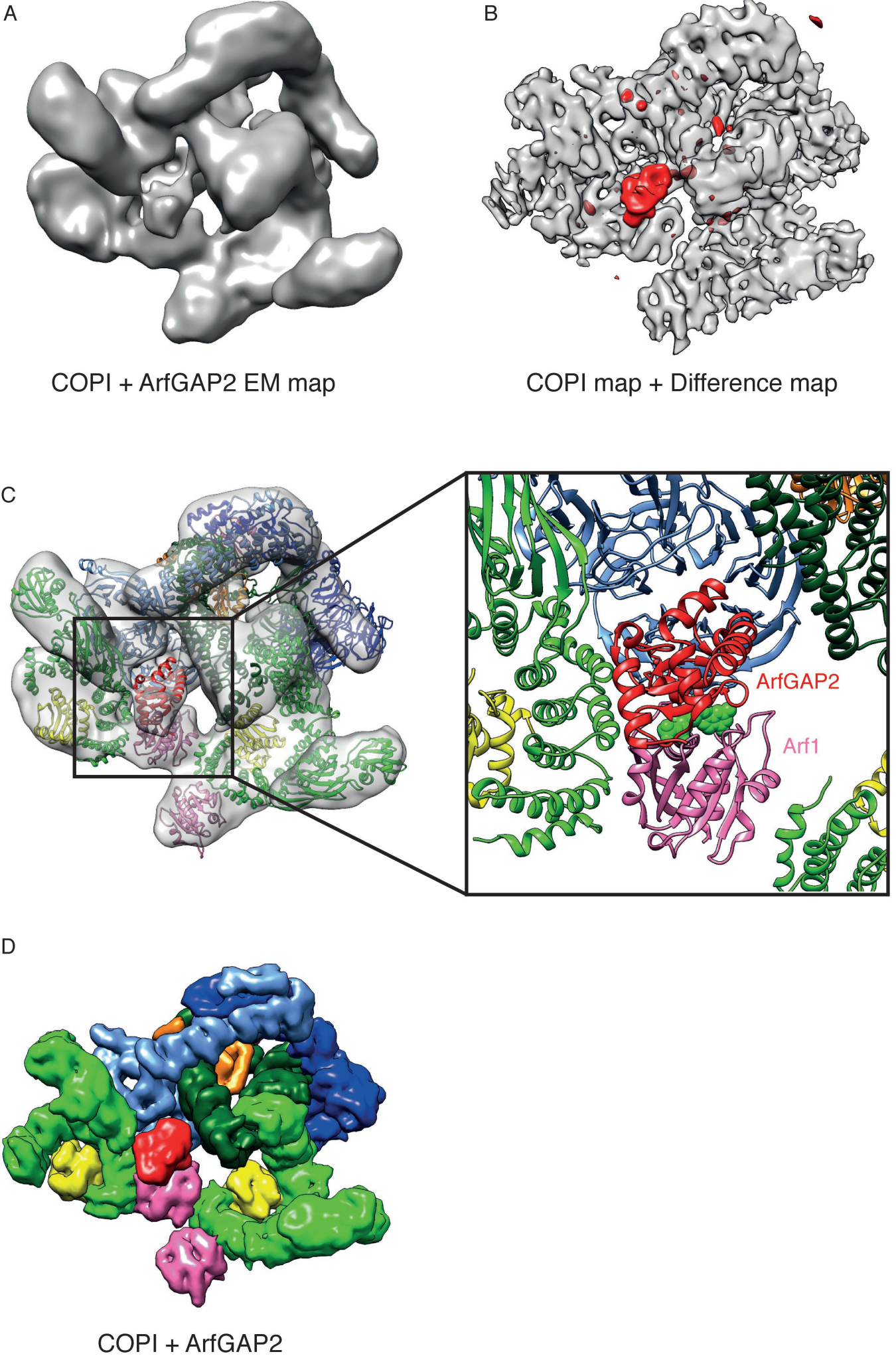
Figure 6





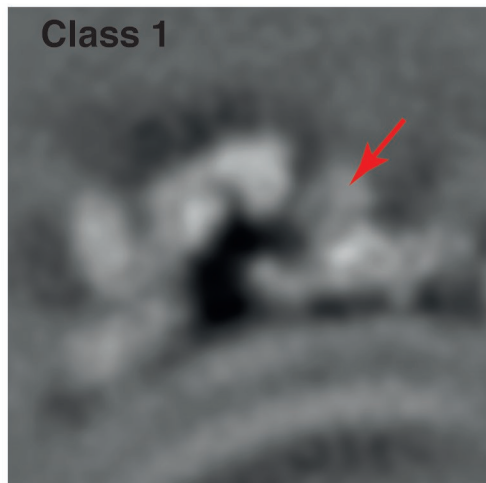




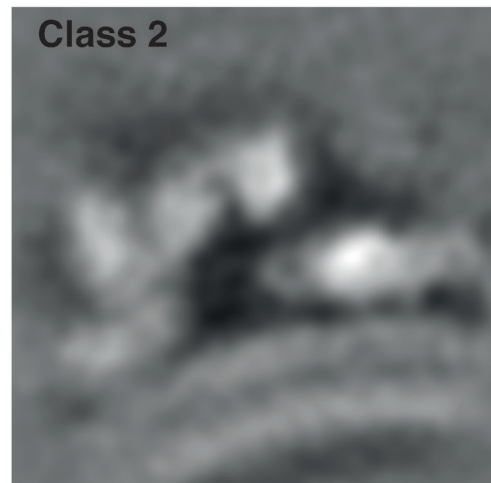




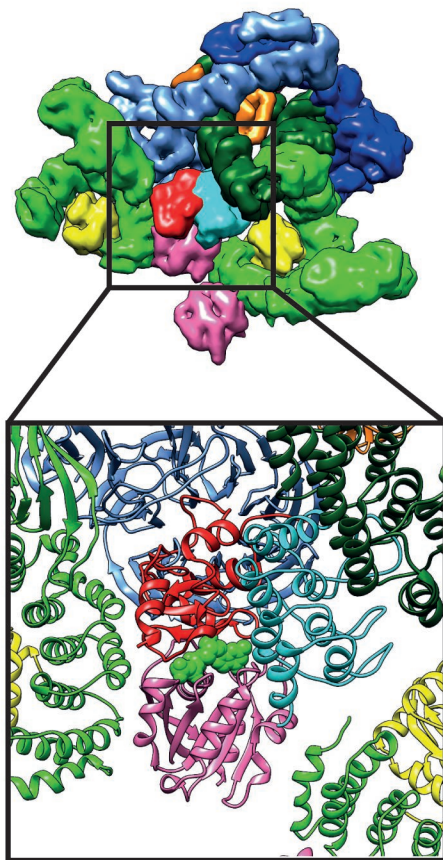
A



B

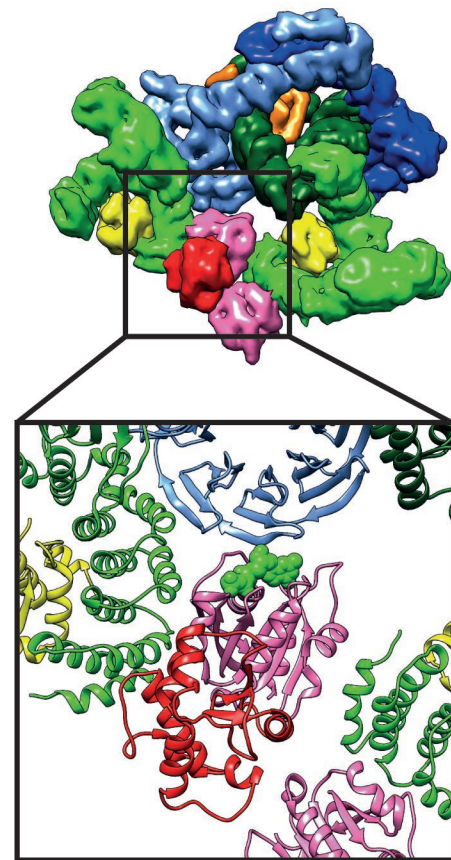


C



**ASAP3: catalytic and  
ankyrin domains  
superimposition**

D



**ArfGAP1  
superimposition**