Structure and Organization of Coat Proteins in the COPII Cage

Stephan Fath,1,2 Joseph D. Mancias,1,2 Xiping Bi,1 and Jonathan Goldberg1,*

1 Howard Hughes Medical Institute and the Structural Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA
2 These authors contributed equally to this work.
*Correspondence: jonathan@ximpact4.ski.mskcc.org
DOI 10.1016/j.cell.2007.05.036

SUMMARY

COPII-coated vesicles export newly synthesized proteins from the endoplasmic reticulum. The COPII coat consists of the Sec23/24-Sar1 assembly unit that can polymerize into an octahedral cage and deform the membrane into a bud. Crystallographic analysis of the assembly unit reveals a 28 nm long rod comprising a central α-solenoid dimer capped by two β-propeller domains at each end. We construct a molecular model of the COPII cage by fitting Sec13/31 crystal structures into a recently determined electron microscopy density map. The vertex geometry involves four copies of the Sec31 β-propeller that converge through their axial ends; there is no interdigitation of assembly units of the kind seen in clathrin cages. We also propose that the assembly unit has a central hinge—an arrangement of interlocked α-solenoids—about which it can bend to adapt to cages of variable curvature.

INTRODUCTION

Vesicle transport pathways are responsible for the propagation and maintenance of organelles, and the underlying mechanisms of vesicle budding and fusion are conserved among eukaryotes (Bock et al., 2001). Budding occurs when cytoplasmic coat proteins assemble on a membrane surface, capture cargo molecules, and polymerize into spherical cages to bud off cargo-laden vesicles (Bonifacino and Glick, 2004; Kirchhausen, 2000). Cells contain a variety of coats—including COPI, COPII, and numerous clathrin/adaptor complexes—with each coat budding vesicles from a discrete subcellular location.

COPII-coated vesicles form on the endoplasmic reticulum (ER) to transport newly synthesized cargo to the Golgi complex. Three proteins—Sec23/24, Sec13/31, and the ARF-family GTPase Sar1—are sufficient to bud 60-nm COPII vesicles from native ER membranes and from synthetic liposomes (Barlowe et al., 1994; Matsuoka et al., 1998). COPII budding is initiated by the activation of Sar1 to its GTP-bound form, causing it to translocate to the membrane and embed an N-terminal α helix in the bilayer (Antonny et al., 1997). Sar1-GTP recruits Sec23/24 to form a Sec23/24-Sar1 “prebudding complex” that binds directly to cargo molecules (Miller et al., 2003; Mossessova et al., 2003). The prebudding complex has a concave surface that conforms to the shape of the underlying membrane vesicle in order to maximize the opportunities for interactions with the membrane and with membrane-proximal elements of cargo proteins (Bi et al., 2002). Finally, the Sec13/31 complex is recruited through an interaction with Sec23/24, and it self-assembles into a polyhedron, leading to membrane deformation (Barlowe et al., 1994; Shaywitz et al., 1997; Stagg et al., 2006).

Electron microscopy (EM) studies of the isolated Sec13/31 assembly unit reveal a 28–50 nm rod, consistent in size with a heterotetramer composed of two Sec13 and two Sec31 polypeptides (Lederkremer et al., 2001; Matsuoka et al., 2001). Recently, the self-assembly reaction has been reconstituted using purified Sec13/31 protein, and the resultant cages have been analyzed by cryo-EM at 30 Å resolution (Stagg et al., 2006). This reveals a striking cuboctahedral architecture for the COPII lattice, built from 24 copies of the Sec13/31 assembly unit. The assembly unit constitutes the edge of the cuboctahedron, and four assembly units converge to form each of 12 vertices (Stagg et al., 2006).

The organization of the COPII cuboctahedron contrasts with that of clathrin cages, which remain the best-characterized vesicular carriers (Kirchhausen, 2000; Pearse et al., 2000). The clathrin lattice is built from triskelion assembly units—trimers of clathrin heavy chain—that are centered on the vertices of the cage, and the 47 nm long “legs” of the clathrin heavy chain interdigitate extensively with neighboring legs as they extend toward the adjacent vertices (Fotin et al., 2004b). Thus, both the symmetry and the nature of the vertex contacts seem to be completely distinct in clathrin and COPII lattices, but the relationship between these vesicular cages remains unknown in the absence of a molecular model of the COPII cuboctahedron.
In this paper, we describe the atomic structure of the Sec13/31 assembly unit, and we present a molecular model of the COPII cage built by fitting the Sec13/31 crystal structures into the cryo-EM density map. We define the geometry of the multivalent contacts between assembly units at the cage vertices, the organizing principle for propagating the COPII cuboctahedron, and we compare the fundamentally different architectures of COPII and clathrin lattices.

RESULTS AND DISCUSSION

Domain Structure of Sec13/31
We prepared the yeast Sec13/31 complex by coexpression in baculovirus-infected insect cells and subjected the purified protein to limit proteolysis (Figure 1A; see also Dokudovskaya et al., 2006). Three protease-resistant products were obtained: the N-terminal .rejected propeller domain of Sec31 formed from seven WD40 sequence repeats or blades; the short C-terminal α-solenoid region of Sec31; and a complex comprising the central α-solenoid domain of Sec31 (residues 370–763) and full-length Sec13, a β-propeller protein containing six WD40 repeats. Absent from the proteolysis products was a stable portion of the proline-rich region, the domain of Sec31 that interacts with Sec23/24 (Shaywitz et al., 1997; Shugrue et al., 1999). Although transiently stable fragments were observed in some proteolysis reactions, we could not identify a robustly stable core, and we tentatively conclude that the proline-rich region is unstructured within the assembly unit (Figure 1A; 20% of residues 770–1110 are proline). Since this implies that the C-terminal α-solenoid domain of Sec31 may be flexibly linked to the remainder of the Sec13/31 complex, we tested whether this domain

Figure 1. Structural Analysis of the Sec13/31 Complex
(A) Diagram shows the domain structure of S. cerevisiae Sec13/31 as defined by proteolytic mapping (see also Dokudovskaya et al., 2006). The products of limit proteolysis reactions were separated chromatographically, and discrete domains were sequenced to identify N termini (c, product of chymotrypsin cleavage; t, trypsin cleavage) and C termini (m, site determined by mass spectrometry measurement of the upstream domain).

(B) Ribbon representation of the Sec13/31 vertex element; the structure was determined by molecular replacement and refined to 3.3 Å resolution. The complex comprises full-length Sec13 colored orange and residues 1–411 of Sec31 colored green (corresponding to the N-terminal region colored black in [A]). β strands are drawn as arrows and α helices as cylinders.

(C) Ribbon diagram of the Sec13/31 edge element; the crystal structure was determined by MAD/SAD phasing and refined to 2.35 Å resolution. This complex comprises full-length Sec13 (red and orange) and the central α-solenoid domain of Sec31 (light and dark green) and forms a Sec13/Sec31-Sec31/Sec13 heterotetramer. The structure is viewed along the molecular dyad axis, oriented as in (B).

(D) Experimental electron density map (2.9 Å resolution, contoured at 2.0 σ) of the crystal asymmetric unit of the edge element, calculated using the combined MAD and SAD phases following density modification. View is rotated 90° from (C).
might fold back to interact with the N-terminal regions, using a protein pull-down assay. We detected no interactions between the C-terminal α-solenoid domain and the β-propeller or central α-solenoid domains of Sec31 or with Sec13; nor did we detect evidence for self-association or interaction with the Sec23/24-Sar1 complex (data not shown).

This preliminary analysis suggested that the C-terminal portion of Sec31, encompassing the proline-rich region and terminal α-solenoid domain (residues 764–1273), might lie outside the “architectural core” of the assembly unit. Indeed, the crystallographic analysis described below confirms that the 28 nm long core particle is a Sec13/31–Sec13/Sec13 heterotetramer constructed from Sec13 plus the N-terminal region of Sec31 (the β-propeller and central α-solenoid domains).

**Crystal Structure Determination**

The atomic model of the assembly unit was built from two crystal structures of yeast Sec13/31 subcomplexes that we refer to as the edge and vertex elements (Table 1). Crystals of the edge element, comprising full-length Sec13 β-propeller and residues 370–763 (central α-solenoid of Sec31, grew in space group P2₁ and contain a Sec13/31–Sec13/Sec13 heterotrimer in the asymmetric unit (Figure 1C). The initial electron density map was calculated to 2.9 Å resolution using combined phases from multi- and single-wavelength anomalous diffraction (MAD and SAD) experiments with selenium as the anomalous scatterer (Figure 1D shows the density-modified map). The structure of the edge element was refined to 2.35 Å resolution (Table 1). In order to facilitate this structure determination we mutated multiple conserved leucine residues to methionine (see Experimental Procedures). For completeness, we crystallized and refined a structure of the native protein to 2.45 Å resolution (Table 1). There are no significant conformational differences between selenomethionine mutant and native structures.

**Table 1. Data Collection and Refinement Statistics**

<table>
<thead>
<tr>
<th>Sec13/31 Complex</th>
<th>Edge</th>
<th>Edge</th>
<th>Edge</th>
<th>Vertex element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set</td>
<td>MAD</td>
<td>SAD</td>
<td>Native</td>
<td>Native</td>
</tr>
<tr>
<td>PDB Accession #</td>
<td>2PM7</td>
<td>2PM6</td>
<td>2PM9</td>
<td></td>
</tr>
<tr>
<td>Space Group:</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P₂₃</td>
<td></td>
</tr>
<tr>
<td>Cell Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å) / (°)</td>
<td>128.1, 53.2, 133.1 108.4</td>
<td>128.2, 52.5, 133.1 108.3</td>
<td>128.4, 52.3, 133.1 108.6</td>
<td>155.2, 155.2, 59.9</td>
</tr>
<tr>
<td>Data Processing</td>
<td>Peak</td>
<td>Inflection</td>
<td>Remote</td>
<td>Peak</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9795</td>
<td>0.9798</td>
<td>0.9643</td>
<td>0.9795</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50–2.8</td>
<td>50–2.8</td>
<td>50–2.8</td>
<td>50–2.8</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>6.2 (29.0)</td>
<td>5.7 (31.8)</td>
<td>6.3 (40.2)</td>
<td>4.7 (24.7)</td>
</tr>
<tr>
<td>R =</td>
<td>25.8 (5.1)</td>
<td>25.3 (4.4)</td>
<td>23.0 (3.4)</td>
<td>26.7 (4.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.6 (99.9)</td>
<td>99.7 (99.6)</td>
<td>99.8 (99.4)</td>
<td>97.1 (91.1)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.6 (3.7)</td>
<td>3.7 (3.6)</td>
<td>3.7 (3.6)</td>
<td>3.4 (3.2)</td>
</tr>
</tbody>
</table>

Refinement Statistics

| Data Range (Å) | 30–2.35 | 30–2.5 | 40–3.3 |
| Reflections    | 68541 | 54986 | 18996 |
| Nonhydrogen Atoms | 9938 | 9938 | 5170 |
| Water Molecules | 261 | 261 | 28 |
| r.m.s. Δ Bonds (Å) | 0.006 | 0.007 | 0.008 |
| r.m.s. Δ Angles (°) | 1.2 | 1.3 | 1.6 |
| R Factor (%) | 24.2 | 25.2 | 25.1 |
| Rfree (%) | 29.8 | 30.1 | 30.5 |

* Rmerge = 100 * Σₙᵢ | lᵢ(h) – <l(h)> | / Σₙᵢ<l(h)>, where lᵢ(h) is the ith measurement and <l(h)> is the weighted mean of all measurement of l(h) for Miller indices h.

* Root-mean-squared deviation (rms Δ) from target geometries.

* R factor = 100 * Σᵢ | Fᵢ(calc) – Fᵢ(calc)/Σᵢ Fᵢ(calc) |

* Rfree was calculated with 5% of the data.

* Highest resolution shell is shown in parentheses.
and refined with data to 3.3 Å resolution (Figure 1B). The polypeptide tracing was assisted by the regular arrangement of bulky tryptophan (or phenylalanine) residues of the WD signature on the A- and C-strands of the Sec31 β-propeller domain (Figure S1).

The crystal structures share a full-length Sec13 β-propeller, and since this domain shows negligible conformational differences in the two crystal forms, it was possible to combine the structures straightforwardly to construct a composite model for the assembly unit, the 28 nm long Sec13/31-Sec31/Sec13 heterotetramer (Figures 2A and 2B).

Structure of the Assembly Unit
The structure of the COPII assembly unit comprises a central α-solenoid dimer capped by two β-propeller domains at each end (Figures 1B, 1C, 2A, and 2B). The α-solenoid regions of the two Sec31 molecules interact about a 2-fold symmetry axis to form a 14 nm long central rod (Figure 1C). The rod is relatively straight and is uniformly ~30 Å in diameter since it is formed from two α-solenoids along its length. However, the two Sec31 molecules are not arranged as an extended antiparallel dimer; instead each Sec31 polypeptide folds back on itself, creating an interlocked dimer. Helices α1–α4 fold back to form intermolecular interactions with helices α11–α18, and helices α5–α10 form isologous interactions with the opposite Sec31 molecule (Figure 1C).

The 28 nm long assembly unit is not uniformly straight along its length because the Sec31 and Sec13 β-propellers are displaced from the axis of the α-solenoid

Figure 2. Organization of the Assembly Unit in the COPII Cage
(A) Comparison of the molecular model of the Sec13/31 assembly unit with the asymmetric unit of the cryo-EM map of the mammalian COPII cage (Stagg et al., 2006). The objects are viewed along the local 2-fold rotation axis. The model, shown in space-filling representation, is a composite of the two crystal structures (oriented and colored as in Figures 1B and 1C). The arrows indicate the ~15 Å displacement of the Sec13 β-propellers from the axis of the α-solenoid rod and the corresponding features in the cryo-EM map.

(B) Orthogonal view shows the difference in the angle at the center of the assembly unit. Here, the arrows show the 15–20 Å displacement of the Sec31 β-propellers from the α-solenoid axis.

(C) The molecular model of the heterotetrameric assembly unit was separated into two Sec13/31 heterodimers, and these were fitted independently as rigid bodies into the cryo-EM map (see Experimental Procedures). The picture shows a complete vertex (two asymmetric units of the cage) and is viewed along the 2-fold symmetry axis that runs through the vertex. One symmetry-related pair (colored dark green and orange) converges at the vertex and is labeled proximal; the other symmetry-related pair (light green and red) is labeled distal.

(D) The molecular model of the cage comprises 24 copies of the assembly unit with octahedral or 432 symmetry. Superimposed is the 30 Å cryo-EM density map from Stagg et al. (2006).
rod (Figures 2A and 2B). Specifically, the Sec13 (inner) β-propellers are displaced ~15 Å from the α-solenoid axis, as can be seen most clearly in Figure 2A. Likewise, the Sec31 (outer) β-propellers are displaced 15–20 Å in a direction roughly parallel with the 2-fold axis of the assembly unit (Figure 2B). The off-axis arrangement of the Sec13 and Sec31 β-propeller domains gives the assembly unit its characteristic shape that can be recognized in the cryo-EM density map (Figures 2 and 3).

As the crystallographic results indicate, the Sec31 β-propeller domain is positioned at the ends of the assembly unit to mediate the vertex contacts that propagate the COPII cage (Figures 1–4). This is surprising from a structural perspective, as it means that the Sec13 molecule separates the β-propeller and α-solenoid domains of Sec31. Figures 4B and 4C show how the Sec31 polypeptide passes directly through Sec13 and contributes a seventh blade (Sec31 residues 380–406) to the six blades of the Sec13 β-propeller. This intimate interaction between the polypeptide chains is reinforced by extensive interfaces involving both axial ends of the Sec13 β-propeller: one end of Sec13 interacts with the Sec31 β-propeller (Figure S2); the other end interacts with the Sec31 α-solenoid domain (via the α14–α15 and α16–α17 loops plus helix α18 of Sec31; see Figure 1C). As a result, the β-propellers probably maintain a fixed orientation relative to the α-solenoid rod. This is important because it determines the geometry and hence the contact interfaces through which the Sec31 β-propellers converge at the vertex.

**Fitting Sec13/31 into the Cryo-EM Density Map**

The dimensions of the composite model of the Sec13/31 heterotetramer (Figures 2A and 2B) are consistent with images of the isolated assembly unit obtained in two EM studies, one of which employed negative staining and the other rotary shadowing to visualize COPII proteins (Lederkremer et al., 2001; Matsuoka et al., 2001). However, both studies revealed a range of conformations for the assembly unit in EM images that involved variation in the angle between the terminal domains and the center of the rod. In both studies it was inferred that the 28 nm long rod may be flexible about a central “protein hinge” (Lederkremer et al., 2001; Matsuoka et al., 2001).

The comparison of the molecular structure and the cryo-EM map of Stagg et al. (2006) provide additional evidence for a hinge located at the interlocked α-solenoid center of the rod (Figures 2A and 2B). When viewed along the 2-fold axis, the overall shape of the molecular model corresponds closely to the asymmetric unit of the cuboctahedron (Figure 2A). However, the orthogonal view (Figure 2B) reveals a major difference between the two structures: in the cryo-EM map, the two halves (two Sec31 molecules)
of the α-solenoid rod are bent 45° from parallel about the center of the rod, whereas in the crystal structure the angle is just 15°. This evidence for a hinge, together with the earlier EM analyses, suggests a mechanism by which the Sec13/31 assembly unit could adapt to lattices of variable curvature, as discussed in more detail below. (Hinge is not meant to imply that the α-solenoid interface is disordered in the crystal structure of the edge element; this region is in fact well ordered with lower crystallographic B-factors than the average for the molecule; Figure S3).

For fitting into the cryo-EM map, we separated the molecular model of the assembly unit into two halves (at the α-solenoid interface) and fitted these independently as rigid bodies into the 30 Å map (Figures 2C and 2D). Initial fitting was done manually and then improved via real- and reciprocal-space approaches, which gave essentially the same result from a range of refinement starting points (see Experimental Procedures). The complete molecular model of the COPII cage, comprising 24 copies of the assembly unit with octahedral symmetry, is illustrated in Figure 2D. The quality of the fit is indicated in Figure 3, which shows the correspondence of the Sec13 and Sec31 β-propellers to features in the cryo-EM map (the map correlation coefficient for the fit shown in Figures 2 and 3 is 0.85). We did not attempt to model in detail the hinge region at the interlocked Sec31-Sec31 center owing to the low resolution of the cryo-EM data.

**Vertex Geometry and Molecular Model for the COPII Cage**

The Sec13/31-Sec31/Sec13 assembly unit constitutes the asymmetric unit of the COPII cuboctahedron (Stagg et al., 2006). This means that although the assembly unit is a symmetric dimer, it is utilized as an asymmetric rod to construct the cage (in other words, the 2-fold symmetry operator of the assembly unit does not belong to the 432 point group of the cage). Importantly, this also means that the two ends of the assembly unit are in different environments at the cage vertices. At one end, a Sec31 β-propeller converges at the vertex dyad axis and interacts with its β-propeller symmetry partner—we refer to this as the proximal Sec31 β-propeller (Figures 2C, 3, and 4A). At the other end of the assembly unit, the distal β-propeller resides ~20 Å from the vertex dyad axis and does not interact with its symmetry partner.

As this description implies, the vertex is constructed from four Sec31 β-propellers: a proximal pair and a distal pair. Figure 3 illustrates the excellent fit of the proximal and distal β-propellers to the cryo-EM density. Since both ends of the assembly unit were fitted into the map as rigid bodies, this suggests that the disposition of the α-solenoid and β-propeller domains is essentially the same at the proximal and distal ends. We surmise that flexibility is limited to the interlocked α-solenoid center of the rod and that plasticity at the β-propeller interdomain contacts is not required for the optimal positioning of the proximal and distal ends in their distinct environments at the vertex. This proposal is substantiated by the aforementioned extensive interfaces between the β-propeller and α-solenoid domains.

When viewed along its 2-fold axis, as in Figure 2C, the vertex can be described as a distorted 4-fold center such that the proximal and distal β-propellers converge on the vertex in a similar orientation. As a consequence, all four Sec31 β-propellers are positioned with their flat axial ends mediating the major contact interfaces (Figure 4A). This geometry is imposed by the relative juxtaposition of the β-propeller and α-solenoid domains along the rod. Specifically, the axis of the Sec13 β-propeller is parallel with, and displaced ~15 Å from, the α-solenoid axis; the
Sec31 β-propeller axis in turn is inclined at a 50° angle to the Sec13 β-propeller axis (Figure 4).

Figure 4A illustrates the relationship between the pairs of proximal and distal Sec31 β-propellers. The assembly unit is pentavalent, forming three contact interfaces via its proximal end (labeled cl, cII, and cIII in Figure 4A) and two at the distal end (cII and cIII). Thus, the molecular model of the cage comprises 24 pentavalent assembly units that form 12 vertices. The multivalent nature of the vertex interactions is the principle for organizing a tetramer at the vertex, and the individual pairwise contacts (cl, cII, and cIII) and the assembly-unit dimers that correspond to them are likely to be exceedingly weak. The cl proximal-proximal contact appears not to involve the entire surface of the axial end of the Sec31 β-propeller because the β-propeller is tilted ~13° from the vertical 2-fold axis. This slight cant of the proximal β-propellers is suggested by a depression in the cryo-EM map at the vertex center (Figure 3B). We have not explored other features of the vertex contacts in atomic detail at this stage because the current cryo-EM data provide for a relatively low-resolution model of the cage.

For descriptive purposes we referred to the vertex as a distorted 4-fold center. In reality the distal β-propellers are translated ~15 Å in toward the center of the cage relative to the proximal β-propellers (Figure 3C). Thus, the clII proximal-distal interface joins two Sec31 β-propellers that are related by an ~90° rotation and an ~15 Å translation (Figure 4A). A similar interaction to clII is observed as a crystal contact in the crystals of the vertex element, where it propagates the 43 crystallographic screw axis (c = 59.9 Å). Our model predicts that the vertex element illustrated in Figure 1B should form a tetramer in solution at sufficiently high concentrations. However, we detected only monomeric Sec13/31(1–411) in gel filtration experiments using protein concentrations up to 30 mg/ml (data not shown). We propose that in the COPII polyhedron, as in viral capsids, the ligation of multiple very weak vertex contacts yields a globally stable cage. Weak vertex interactions are presumably important in the self-assembly reaction to prevent the accumulation of too many nuclei—we assume that Sec13/31 nucleation involves the low-probability association of four assembly units to form a vertex. The other general possibility in this context is that assembly is assisted by a template protein—an assembly chaperone—though there is no evidence for such a factor in COPII coat formation.

In summary, the positions of the β-propeller and x-solenoid domains in the COPII cage have been assigned unambiguously to structural features of the 30 Å cryo-EM map. Although the details of the vertex contacts in our model of the cage cannot be assessed at this resolution, it is nevertheless clear that the pairwise contacts of the Sec31 β-propellers look much like conventional protein-protein interactions and that the vertex is not constructed through the interdigitation of assembly units. Thus, both the symmetry and nature of the vertex contacts are completely distinct in COPII and clathrin cages.

Finally, Stagg et al. (2006) observed that the 2-fold symmetry axis of the assembly unit is not located centrally along the edge of the cuboctahedron. Consistent with this, we find that the 30 Å diameter central x-solenoid dimer fits well into the cryo-EM map and the interlocked Sec31-Sec31 center coincides with the 2-fold symmetry center of the assembly unit in the map (Figures 2 and 3). Although the symmetry of the assembly unit is not connected to the symmetry of the cage, the assembly unit is oriented such that its 2-fold axis passes close to the center of the cuboctahedron (Figure 3C).

COPII and Clathrin Lattices

COPII and clathrin cages are both constructed from x-solenoid and β-propeller building blocks (Fotin et al., 2004b; ter Haar et al., 1998; Ybe et al., 1999). We observe a close correspondence of the COPII and clathrin x-solenoid folds despite some divergence at the point where the two Sec31 chains interlock (in the vicinity of helix 2α; see Figure S4). In particular, both x-solenoids follow a straight path that is distinct from the highly curved forms observed in other x-solenoid structures (Kajava, 2002). (The x-solenoid fold is also termed α-helical zigzag; ter Haar et al., 1998.)

x-solenoid and β-propeller folds have also been identified through sequence analysis in COPI coat proteins and in the yNup84/vNup107–160 subassembly of the nuclear pore complex. On this basis it has been posited that a common ancestor of vesicle-coat and nuclear-pore complexes—a membrane-curving module—arose with the emergence of the endomembrane system to stabilize curved membranes (Devos et al., 2004). It is surprising then that the x-solenoid and β-propeller domains are arranged in such fundamentally different ways in COPII and clathrin such that they form distinct symmetry centers and assembly-unit interactions that propagate the cages. In clathrin, a triskelion assembly unit lies at each vertex, and the x-solenoid legs of neighboring triskelia interdigitate extensively as they extend toward the adjacent vertices; the β-propeller is not part of the architectural core and instead projects in toward membrane to interact with adaptor molecules (Fotin et al., 2004b; Kirchhausen, 2000). In contrast, the COPII assembly unit is a rod that constitutes the edge of a cuboctahedron, and four rods converge to form the vertex with no interdigitation of assembly units. x-solenoid domains form the core of the edge, but, unlike clathrin, the COPII vertices are formed from β-propellers. In summary, the COPII and clathrin lattices seem not to share common construction principles other than the use of x-solenoid and β-propeller folds. Future structural studies of the COPI cage and the nuclear pore complex should reveal whether common principles exist and will shed more light on the evolutionary relationships among these membrane-curving proteins.

A Possible Mechanism for Lattice Adaptability

The ability of COPII coat proteins to control ER exit of very large cargo molecules such as procollagen and...
Figure 5. α-Solenoid Crossing Angle and a Possible Mechanism for Lattice Adaptability
(A) COPII. Comparison of the crystal structure with the cryo-EM map, focused on the crossing angle of the Sec31 α-solenoids. The Cα trace shows the dimerization region, with the Sec31 polypeptides colored light and dark green. The molecular dyad axis is vertical. The corresponding view of the cryo-EM map suggests a 30° rotation between the two Sec31 α-solenoid rods relative to the crystal structure (also see Figure 2B). The symbols indicate the vertices with their dyad symmetry; the hinge is orientated such that a decrease in the hinge angle would yield a smaller cage, according to our model.

(B) Clathrin. Comparison of molecular models of the “proximal pair” of α-solenoids based on cryo-EM analysis of the clathrin mini-coat (top) and hexagonal barrel (bottom; Fotin et al., 2004b); the models differ by an 8° rotation. The vertices have 3-fold rotational symmetry, and the hinge is inverted relative to COPII such that a decrease in the hinge angle yields a larger cage. We do not mean to imply that the overlapped α-solenoids interact in a similar manner in COPII and clathrin; indeed, only the COPII α-solenoids form intimate contacts across the interface.

chylomicron raises the question of whether COPII vesicles larger than the standard 60 nm cages can form from lattices of altered curvature (Fromme and Schekman, 2005). Larger (>60 nm) cages have been identified in cryo-EM images of self-assembled Sec13/31 but have not as yet been analyzed in detail (Stagg et al., 2006). Lattice adaptability of this kind has been observed for clathrin, whereby cargo size can influence triskelion self-assembly and dictate the dimensions of the cage (Crowther et al., 1976; Fotin et al., 2004b; Heuser, 1980; Ehrlich et al., 2004). At the molecular level, a larger clathrin cage that incorporates additional assembly units requires changes in lattice contacts, and this is achieved through an alteration of the crossing angle between adjacent α-solenoids along the edge (Fotin et al., 2004b; Figure 5B). We propose that a similar mechanism may control the size of the COPII cage (Figure 5). That is, we envisage a larger COPII cage that incorporates additional assembly units (with additional polygonal facets) and a concomitant alteration of the crossing angle at the Sec31-Sec31 protein hinge. At the center of the assembly unit, the putative protein hinge involves two Sec31 α-solenoid chains that overlap to form a 50 Å long interface (Figures 1C and 5), an arrangement that resembles the overlapping clathrin proximal and distal α-solenoid pairs (though only the COPII α-solenoids form intimate contacts across the interface). If flexibility at the α-solenoid crossing angle compromises cage integrity, in clathrin this might be counteracted by the extensive interdigitation of triskelia, whereas in COPII the interlocked arrangement of Sec31 α-solenoid chains may be an important stabilizing factor (Figure 1C).

The notion that the Sec13/31 assembly unit is flexible about a central protein hinge was proposed based on two EM studies (Lederkremer et al., 2001; Matsuoka et al., 2001) and is reinforced by the observed difference in hinge angle between the X-ray structure and cryo-EM map (Stagg et al., 2006). The protein-hinge model and its relation to the clathrin system will be tested as cryo-EM analyses of the COPII cuboctahedron and larger cages are extended to higher resolution.

Functional Organization of the COPII Coat
The model in Figure 6 illustrates the arrangement of protein components and membrane in the COPII vesicle. Twenty-four copies of the 28 nm assembly unit organize themselves as a cuboctahedron with an outer diameter of ~60 nm and an inner diameter of ~52 nm (these are average measurements as different parts of the assembly unit reside at different radial distances). For comparison, COPII-coated vesicles synthesized in vitro from ER membranes measure 60–65 nm diameter and contain a 10 nm thick coat that is thought to comprise the ~5 nm cage or “outer layer” plus a 5 nm inner layer of multiple Sec23/24-Sar1 complexes (Barlowe et al., 1994; Lee et al., 2004; Matsuoka et al., 2001). On this basis we modeled the membrane vesicle as a 40 nm sphere, two-thirds of the diameter of the cage, to allow an ~5 nm space for Sec23/24-Sar1, two copies of which are drawn in Figure 6. This estimate of the vesicle diameter may be on the high side, as a recent study of clathrin-coated vesicles employing electron cryo-tomography of individual particles reveals somewhat smaller membrane vesicles, which are on average ≤50% of the diameter of the surrounding protein coat (Cheng et al., 2007).

The Sec13/31 cuboctahedron contains binding sites for 48 copies of the inner-layer Sec23/24-Sar1 complex (only two copies are drawn in Figure 6 for clarity). This assumes that the proline-rich region of Sec31, known to interact with
Sec23/24 (Shaywitz et al., 1997; Shugrue et al., 1999), is functionally accessible on the 24 proximal and 24 distal ends of the assembly unit (the 340 residue proline-rich region follows helix \( \alpha_18 \); see Figure 1C). Since the concave membrane-apposing surface area of Sec23/24-Sar1 measures \( \approx 8,200 \) \( \AA^2 \) (Bi et al., 2002), 48 copies of the complex would cover 75%–80% of the surface area of a 40 nm vesicle. This could in principle leave sufficient space for the cytoplasmic regions of transmembrane cargo proteins—a quantitative analysis of the composition of synaptic vesicles shows that protein transmembrane domains comprise \( \approx 18% \) of the surface area of the outer leaflet (Takamori et al., 2006). However, it seems unlikely that the inner-layer complexes are this densely packed and more reasonable to assume that substoichiometric amounts of Sec23/24-Sar1 are incorporated into the COPII coat. Single-particle analysis of clathrin-coated cargo proteins—a quantitative analysis of the composition of synaptic vesicles shows that protein transmembrane domains comprise \( \approx 18% \) of the surface area of the outer leaflet (Takamori et al., 2006). However, it seems unlikely that the inner-layer complexes are this densely packed and more reasonable to assume that substoichiometric amounts of Sec23/24-Sar1 are incorporated into the COPII coat. Single-particle analysis of clathrin-coated vesicles shows that the inner layer is not as densely packed as would be predicted if AP2 adaptors (or their functional equivalents) were bound stoichiometrically to the clathrin heavy chain (Cheng et al., 2007). And this observation was interpreted in terms of a budding reaction in which the connection of clathrin assembly units to the membrane via inner-layer adaptors is required for the initiation, but not for the completion, of budding (Cheng et al., 2007).

The concept of a vesicular coat as a mechanical device for budding vesicles implies that the driving energy for self-assembly of the spherical cage—which arises from the compulsion to maximize the number of stable bonds between assembly units—is directly coupled to membrane deformation. In the case of the Sec13/31 cage, the connection to lipid membrane is through Sec23/24 to the N-terminal amphipathic \( \alpha \) helix of the Sar1-GTP molecule (Barlowe et al., 1994; Shaywitz et al., 1997). According to the calculations above, as many as 48 (but probably fewer) copies of Sar1-GTP will provide the linkages by which the Sec13/31 assembly units incrementally deform the membrane into an \( \approx 40 \) nm bud. Sar1-GTP may, in addition, play a direct role to initiate membrane curvature since it embeds its amphipathic \( \alpha \) helix in the cytosolic leaflet of the membrane (Bielli et al., 2005; Lee et al., 2005). If we assume a unitary stoichiometry among the COPII components, then the amphipathic \( \alpha \) helices of 30–48 copies of Sar1-GTP will displace an area of lipid that corresponds to 1.4%–2.4% (±0.4%) of the surface area of the outer leaflet of a 40 nm vesicle. Since the area of this outer leaflet is 55%–60% larger than the inner leaflet, the contribution of Sar1-GTP to membrane curvature is more likely to be catalytic than stoichiometric. Conceivably, supernumerary Sar1-GTP molecules incorporate into the bud to impart curvature, but it is unclear how these would be organized in the absence of a link to the other COPII coat components.

The molecular model of the Sec13/31 assembly unit constructed from the crystallographic analysis seems to account for all the density in the cryo-EM map (Figures 2D and 3). Thus, the architectural elements of the assembly unit comprise Sec13 plus the \( \beta \)-propeller and central \( \alpha \)-solenoid domains (residues 1–763) of Sec31 (Figure 1A). The C-terminal region of Sec31 (residues 764–1273) is not part of the architectural core, and we infer that this region is not observed following image reconstruction from electron cryo-micrographs of multiple cages as it does not conform strictly to the symmetry of the cage (Stagg et al.,...
The 55K C-terminal region comprises the 35K proline-rich segment plus the terminal α-solenoid domain (Figure 1A). The low sequence complexity of the proline-rich segment and its sensitivity to proteolysis suggest that this region is largely unstructured and forms a flexible linker to the terminal α-solenoid and that the 55K region projects in toward the membrane to engage Sec23/24 (via a proline-rich segment; Shaywitz et al., 1997). A flexible connection between the Sec13/31 cage and the Sec23/24-Sar1 complex would afford Sec23/24-Sar1 some mobility on the membrane surface, perhaps facilitating the packaging of a variety of cargo proteins of different shapes and sizes.

Sec13/31 forms a more open lattice than clathrin due to the octahedral symmetry of the cage and the distinct composition of the cage edge. Regarding symmetry, the clathrin lattice forms a wide variety of designs of cages but in general seems to require about 60 edges to build a cage around a 40 nm vesicle (Cheng et al., 2007); by contrast the COPII cuboctahedron uses 24 edges to surround a vesicle of comparable size. Regarding the composition of the edge, in the clathrin cage these are composed of four α-solenoid chains, whereas the COPII cage edge involves just two α-solenoids (Fotin et al., 2004b). The net effect is to create an open COPII lattice: in particular, the square face of the cuboctahedron measures ~27 nm from side-to-side and ~33 nm across the diagonal (Figures 2D and 6). This may be large enough to accommodate the neck of the membrane bud, and if this is a stable arrangement it raises the question of whether the completion of cage assembly can in and of itself drive vesicle fission. Recent studies of yeast and mammalian COPII proteins suggest that cage formation is insufficient and reveal a major role for Sar1-GTP in the fission reaction, possibly involving the amphipathic α helix invading and constricting the bud neck (Antony, 2006; Bielli et al., 2005; Lee et al., 2005).

Finally, the distinct organization of COPII and clathrin cages, in terms of symmetry and the interdigitation of assembly units, has implications for the disassembly reaction. The heavily interdigitated clathrin lattice requires hsc70 and the cochaperone auxilin to effect structural perturbations at the vertex that lead to global uncoating (Braell et al., 1984; Fotin et al., 2004a). Disassembly of the Sec13/31 lattice may be more straightforward since its stability relies on fewer vertices and on less extensive interactions between assembly units and there is no evidence that COPII uncoating is a regulated process.

**EXPERIMENTAL PROCEDURES**

**Protein Production**

The full-length and all truncated forms of yeast Sec13/31 were prepared by coexpression in insect cells infected with engineered baculoviruses (relative molecular mass of Sec13, 33K and Sec31, 139K). Sec31 forms were designed with an N-terminal His6 tag. Insect cells prepared by coexpression in insect cells infected with engineered baculoviruses (relative molecular mass of Sec13, 33K and Sec31, 139K).

For selenomethionine incorporation, insect cells were incubated for 24 hr prior to viral infection at 1.5 million cells/ml in Met/Cys-free medium supplemented with 20 mg/ml L-selenomethionine and 2 mM cysteine (medium A). At 5 hr postinfection, cells were pelleted and resuspended in medium A. Following 24 hr incubation, the selenomethionine concentration was increased to 80 mg/ml and cells were harvested after a further 24 hr incubation. Protein was extracted from cell lysates as before. The extent of selenomethionine incorporation into the core Sec13/31 complex was ~75%, as assessed by mass spectrometry. To increase the selenomethionine content, we mutated six conserved leucine residues to methionine in Sec13 (residues 11, 17, 24, 80, 115, and 222) and made five such changes in Sec31 (residues 449, 536, 615, 622, and 674).

**Crystallization and Structure Determination**

The Sec13/31 edge complex (comprising full-length Sec13 and residues 370–763 of Sec31) was crystallized at 4°C by the hanging-drop method. To 1 μl of 35 mg/ml protein solution was added 1 μl of well solution containing 15% polyethylene glycol (PEG) 4000, 10% dimethyl sulfoxide, 6% dioxane, and 100 mM Tris-HCl (pH 7.5). Hanging drops were seeded with microcrystals after two days equilibration. The crystals are monoclinic, space group P21 (a = 128.2 Å, b = 52.5 Å, c = 133.1 Å, β = 108.34°), and contain (Sec13/31)4 in the asymmetric unit. For diffraction studies, crystals were transferred to 15% PEG 4000, 25% dimethyl sulfoxide, and 100 mM Tris-HCl (pH 7.9) and were flash-frozen in liquid propane.

The edge structure was determined via MAD and SAD experiments. MAD data were collected from a single frozen crystal at beamline X25 of the National Synchrotron Light Source (NSLS). Data were processed with the program HKL2000 (Otwonowski and Minor, 1997), and SAD analysis was done with the program SOLVE (Terwilliger and Berendzen, 1999) using data between 38 and 2.6 Å resolution. Solve found 20 of the 30 selenium atom positions and, following refinement, reported a mean figure of merit (fom) of 0.31. MAD data at three wavelengths from 38 to 2.9 Å resolution were collected (also at beamline X25), and analysis with SOLVE gave a fom of 0.44. The MAD and SAD phases were combined using SIGMAA (CCP4, 1994) to give a fom of 0.55 for data between 38 and 2.9 Å resolution. Density modification with noncrystallographic symmetry averaging was carried out with DM (CCP4, 1994), yielding an initial electron density map of high quality. Model refinement using the program CNS (Brünger, 1990) reduced the R factor to a final value of 24.2% (Rfree = 29.8%) for data from the SAD experiment) between 30.0 and 2.35 Å resolution. The model comprises 9938 protein atoms and 261 water molecules, with just three Ramachandran violations (residue 146 in both copies of Sec13 and residue 497 in one copy of Sec31). Finally, native data from 30 to 2.5 Å resolution were collected on a rotating anode X-ray generator and used to refine a model of the native protein (final R factor = 25.2; Rfree = 30.3). In the models (SAD and native), the following residues been omitted due to weak electron density: 370–372, 470–494, 691–693, and 745–763 in Sec31 and residues 1, 158–169, and 292–297 in Sec13. The X-ray data and refinement statistics are summarized in Table 1.

Crystals of the Sec13/31 vertex element were formed at 22°C by the hanging-drop method. This complex comprises residues 1–411 of Sec31 plus full-length Sec13 that contains the six Leu-to-Met mutations described above. Hanging drops containing 15 mg/ml protein, 8% PEG 3350, and 100 mM trisodium citrate, and 10 mM manganese (II) chloride were equilibrated over well solutions of 16% PEG 3350 plus 200 mM trisodium citrate. The crystals (space group P4321, a = 155.2 Å, b = 278.2 Å, c = 278.2 Å, 292–297 in Sec13. The X-ray data and refinement statistics are summarized in Table 1.
c = 59.9 Å, 75% solvent) contain one vertex element in the asymmetric unit. For diffraction experiments crystals were transferred to well solution containing an additional 25% glycerol and were flash-frozen in liquid propane. X-ray data were collected at beamline 24ID of the Advanced Photon Source (APS) and processed as before. The structure was solved by molecular replacement with the program AMORE (CCP4, 1994) using Sec13 as the search model. Initial electron density was markedly improved using the prime-and-switch function of the program RESOLVE (Terwilliger and Berendzen, 1999), and the complete Sec13/31 vertex element was built iteratively with rounds of refinement using CNS (Brünger, 1998). The final R factor is 25.1% (Rsym of 30.5%) for data between 40.0 and 3.3 Å resolution (Table 1). There are two outliers in a Ramachandran plot of the final model (Ser156 and Phe369 of Sec31). The following residues have been omitted due to weak electron density: Sec13 residues 1, 158–168, and 293–297 and Sec31 residues 1–4 and 411.

### Fitting into the Cryo-EM Density Map

The atomic model of the assembly unit (that is, the composite of the two crystal structures) was separated (at the Sec31-Sec31 dimer interface) into two halves, and these were fitted independently as rigid bodies into the 3 Å cryo-EM map determined by Stagg et al. (2006). Initial fitting was done manually using O (Jones et al., 1991) and was improved via real- and reciprocal-space approaches. For real-space optimization, electron density maps for each half of the atomic model were calculated to the nominal resolution of the cryo-EM map and then fitted to the cryo-EM map using MAVE (and related programs from the Uppsala suite; Köváry and Jones, 1999).

Alternatively, the initial fit was improved by reciprocal-space refinement of the two rigid bodies using CNS (Brünger, 1998), having first transformed the cryo-EM map. We imposed a very soft van der Waals penalty on intermolecular (432 symmetry-related) contacts in CNS to restrict overlap of Sec31 β-propellers at the vertex. This energy term was not enforced at the α-solenoid crossover site, and no attempt was made to model the hinge region, as the resolution limit of the cryo-EM data is insufficient for this at present.

The fit used for the illustrations (Figures 2 and 3) is from a reciprocal-space refinement run. The correlation coefficient of the maps following real-space refinement in this case is 0.85 (calculated using MAVE with a 5 Å protein mask). Note that we fitted the yeast molecular model into the cryo-EM density map. We imposed a very soft van der Waals potential on intermolecular contacts (432 symmetry-related) in CNS to restrict overlap of Sec31 β-propellers at the vertex. This energy term was not enforced at the α-solenoid crossover site, and no attempt was made to model the hinge region, as the resolution limit of the cryo-EM data is insufficient for this at present.

### Supplemental Data

Supplemental Data include four figures and can be found with this article online at [http://www.cell.com/cgi/content/full/129/7/1325/DC1/](http://www.cell.com/cgi/content/full/129/7/1325/DC1/).

### ACKNOWLEDGMENTS

We thank John Walker for assistance with data collection and reduction and Anand Saxena for use of synchrotron facilities at NSLS. This work was supported by grants from the NIH and the Howard Hughes Medical Institute (to J.G.), a fellowship from the Deutsche Forschungsgemeinschaft (to S.F.), and MSTP and NRSA fellowships (to J.D.M.).

Received: March 21, 2007
Revised: April 27, 2007
Accepted: May 11, 2007
Published: June 28, 2007

### REFERENCES


Accession Numbers
The atomic coordinates have been deposited in the Protein Data Bank with the accession codes 2PM6, 2PM7, and 2PM9 (see Table 1).
Insights into COPII Coat Nucleation from the Structure of Sec23•Sar1 Complexed with the Active Fragment of Sec31

Xiping Bi,1 Joseph D. Mancias,1 and Jonathan Goldberg1,*
1Howard Hughes Medical Institute and the Structural Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA
*Correspondence: jonathan@xlimpact14.ski.mskcc.org
DOI 10.1016/j.devcel.2007.10.006

SUMMARY

The COPII vesicular coat forms on the endoplasmic reticulum from Sar1-GTP, Sec23/24 and Sec13/31 protein subunits. Here, we define the interaction between Sec23/24•Sar1 and Sec13/31, involving a 40 residue Sec31 fragment. In the crystal structure of the ternary complex, Sec31 binds as an extended polypeptide across a composite surface of the Sec23 and Sar1-GTP molecules, explaining the stepwise character of Sec23/24•Sar1 and Sec13/31 recruitment to the membrane. The Sec31 fragment stimulates GAP activity of Sec23/24, and a convergence of Sec31 and Sec23 residues at the Sar1 GTPase active site explains how GTP hydrolysis is triggered leading to COPII coat disassembly. The Sec31 active fragment is accommodated in a binding groove supported in part by Sec23 residue Phe380. Substitution of the corresponding residue F382L in human Sec23A causes cranio-lenticulo-sutural dysplasia, and we suggest that this mutation disrupts the nucleation of COPII coat proteins at endoplasmic reticulum exit sites.

INTRODUCTION

Vesicle transport pathways support the growth and maintenance of organelles. Vesicles are formed in a budding reaction involving the coordination of two coat protein complexes: an “inner shell” complex that binds to the membrane and captures cargo molecules, and an “outer shell” complex that polymerizes into a spherical cage to deform the membrane into a bud. A protein-protein link between the inner shell and the self-assembling outer shell—as observed for example between clathrin heavy chain and adaptors—is needed to couple the actions of these protein complexes, leading to coat protein clustering and cargo concentration into a membrane bud (Kirchhausen, 2000; McMahon and Mills, 2004).

COPII-coated vesicles bud from the endoplasmic reticulum (ER) to export newly synthesized proteins to the Golgi complex (Lee et al., 2004). The COPII coat forms through the sequential binding of three cytoplasmic proteins—Sar1, Sec23/24 and Sec13/31—to the ER membrane (Barlowe et al., 1994). Budding is initiated by the activation of the G protein Sar1 to its GTP-bound form, causing it to embed an N-terminal α helix in the bilayer (Antonny et al., 1997). Sar1-GTP recruits Sec23/24 to form the inner shell or “prebudding” complex that binds directly to cargo molecules (Miller et al., 2003; Mossessova et al., 2003). Finally, the prebudding complex recruits Sec13/31 to nucleate the polymerization of multiple Sec13/31 assembly units into an octahedral cage that constitutes the outer shell of the COPII coat (Barlowe et al., 1994; Stagg et al., 2006).

Structural analysis of Sec23/24•Sar1 reveals a bowtie-shaped complex with a concave surface for binding the membrane vesicle; an extensive interface between Sec23 and Sar1 is stabilized by the bound GTP molecule (Bi et al., 2002). The Sec13/31 assembly unit is a heterotetramer—comprising Sec13/Sec31•Sec31/Sec13—the architectural core of which is organized as a linear array of α-solenoid and β-propeller domains to form a 28-nm long rod (Fath et al., 2007; Lederkremer et al., 2001; Matsuoka et al., 2001; Stagg et al., 2006). Twenty-four copies of the rod assemble to form the COPII cuboctahedron (Fath et al., 2007; Stagg et al., 2006).

The nature of the link between Sec23/24•Sar1 and Sec13/31 remains unclear. Studies of yeast and mammalian proteins suggest that an ~300 residue proline-rich region of Sec31 interacts with Sec23 (Shaywitz et al., 1997; Shugrue et al., 1999). This region is not part of the architectural core of the Sec31 assembly unit, and it seems not to fold into a structured domain (Fath et al., 2007). On this basis it has been suggested that the proline-rich region forms a somewhat flexible linker that projects from the outer shell cage in toward the membrane to engage Sec23/24 (Fath et al., 2007).

Defining the interactions between the inner and outer shell complexes is important not only for understanding COPII assembly, but also coat disassembly triggered by GTP hydrolysis. Rapid GTP hydrolysis on Sar1 requires Sec23/24, which is the GAP (GTPase-activating protein) for the reaction (Yoshihisa et al., 1993), and is accelerated an additional order of magnitude by Sec13/31 (Antonny et al., 2001). Thus, the GTP hydrolysis reaction on Sar1 is programmed in the COPII budding process, so as to couple coat assembly to disassembly.
In this paper, we define the active fragment of Sec31 that binds and stimulates the GAP activity of Sec23/24. We describe the 2.5 Å crystal structure of the ternary complex comprising the active fragment bound to Sec23-Sar1 stabilized with a nonhydrolyzable GTP analog. The structure reveals that Sec31 binds as an extended polypeptide across the Sec23 and Sar1 molecules and inserts amino-acid side chains in the vicinity of the Sar1 active site to accelerate GTP hydrolysis in combination with Sec23.

The F382L substitution in human Sec23A that causes cranio-lenticulo-sutural dysplasia maps close to the binding site for the Sec31 active fragment, implying that the mutation may disrupt nucleation of the COPII protein machinery. The functional consequences of the disease mutation are revealed in this issue of Developmental Cell in the accompanying manuscript by Fromme et al. (2007).

RESULTS

**Sec13/31 Stimulates the GAP Activity of Sec23/24 in Solution**

The ability of Sec13/31 to accelerate GAP activity of Sec23/24 was discovered using S. cerevisiae COPII proteins bound to synthetic liposomes (Antony et al., 2001). According to the results of that study, there are two general mechanisms by which Sec13/31 might alter the Sar1 active site to stimulate GAP activity: either a simple mechanism in which an element of Sec13/31 forms a stoichiometric contact with Sec23/24 or Sar1, or a more complex mechanism in which the polymerization of the coat induces changes at the active site—this might be triggered by intersubunit contacts, perhaps in response to membrane curvature.

We favored the simple mechanism, which predicts that in vitro GAP stimulation should be retained in the absence of phospholipid membrane and should require only a portion of the Sec13/31 molecule. We used the truncated, soluble form of S. cerevisiae Sar1 (residues 24–190, hereafter referred to as Sar1) in a fluorometric assay, and tested first whether GAP stimulation could be detected using full-length COPII subunits dispersed in solution. In preliminary experiments we used a substrate comprising Sar1 bound to the fluorescent GTP analog, mant-deoxyGTP (3′-O-[N-methyl-anthraniloyl]-deoxyGTP), because we were concerned that the fluorophore moiety attached to the 2′ position might interfere with the interaction between Sar1 and Sec23 residues that approach close to the ribose 2′-hydroxyl group (Bi et al., 2002). However, GTP hydrolysis experiments using Sar1 bound to mant-deoxyGTP and mant-GTP gave essentially identical results (data not shown), so we used mant-GTP in subsequent experiments (mant-GTP is a mixture of forms with the fluorescent mant moiety attached to either the 2′ or 3′ position).

Figure 1B shows a representative set of experiments in which the fluorescence output of a 1 μM solution of Sar1-mant-GTP was continuously monitored (λ = 438 nm) as COPII proteins were added. In curve (ii), the addition of Sec23/24 alone to the fluorescent Sar1 substrate caused a slow decrease in fluorescence output upon GTP hydrolysis, because mant-GDP has a decreased fluorescence increase in fluorescence intensity upon formation of the Sec23/24•Sar1-mantGTP complex. Curve (i) is a control experiment using mant-GppNHP. Curve (ii) shows GTP hydrolysis by Sec23/24 in the absence of Sec13/31. Curve (iii) shows the additional rate acceleration caused by 4 μM full-length Sec13/31. Curve (iv) shows rapid GTP hydrolysis by Sec23/24 plus 10 μM Sec31 active fragment (residues 899–947). The inset graph shows a least-squares fit to a first order exponential (red line), using data (black line) from an experiment containing Sec23/24 and 40 μM Sec31 active fragment.

**Developmental Cell**

Structure of Sec23•Sar1 Complexed with Sec31
mimics earlier reports of Sec23-dependent GTP hydrolysis on Sar1 (Antonny et al., 2001; Yoshihisa et al., 1993). To confirm that the assay monitors GTP hydrolysis, we bound Sar1 to the nonhydrolyzable form of the fluorescent nucleotide, mant-GppNHp, and found that the addition of Sec23/24 caused no change in fluorescence (Figure 1B, curve i). We then tested the effect of full-length Sec13/31. The addition of 4 μM Sec13/31 to the solution containing Sec23/24 and Sar1-mant-GTP (curve iii) stimulated GTP hydrolysis beyond the rate due to Sec23/24 alone. Sec13/31 itself did not have GAP activity. It synergized with Sec23/24 to accelerate GTP hydrolysis on Sar1, as reported by Antonny et al. (2001).

Sec13/31 may stimulate GAP activity by increasing the intrinsic GTPase reaction rate of the Sec23/24•Sar1 complex or by increasing the affinity of Sec23/24 for Sar1. To delineate these two catalytic contributions, we measured GTP hydrolysis rates across a range of concentrations of Sec23/24, in the presence or absence of 3 μM Sec13/31 (Figure 1C). Reaction rates were obtained by fitting the time course data to a simple exponential function. In all the experiments there is a rapid increase in fluorescence upon formation of the Sec23/24•Sar1 complex (much like the fluorescent transient observed when neurofibromin binds to H-Ras-mant-GTP [Ahmadian et al., 1997]). The subsequent fluorescence decrease, which monitors GTP hydrolysis, fits well to the exponential function (inset in Figure 1B). The results of these preliminary experiments show that the presence of full-length Sec13/31 increases the GTPase rate roughly 4-fold, and increases the affinity between Sec23/24 and Sar1 roughly 2-fold (Figure 1C).

Thus, GAP stimulation by Sec13/31 does not depend on the accretion of COPI coat subunits on a phospholipid membrane. Stimulation can be measured in solution using the mant fluorometric assay and soluble coat subunits, and this provided a straightforward system to dissect the active fragment of Sec13/31.

### A 50 Residue Fragment of Sec31 Binds to Sec23/24•Sar1 and Stimulates GAP Activity

The interaction between Sec13/31 and Sec23 maps to the proline-rich region of the Sec31 polypeptide (Figure 1A shows the domain structure of Sec31). Specifically, the S. cerevisiae Sec31 fragment comprising residues 850–1175 was shown to interact with full-length Sec23 in a two-hybrid analysis (Shaywitz et al., 1997). On this basis we tested whether the proline-rich region could stimulate GAP activity in the same manner as full-length Sec13/31. The proline-rich region has 20% proline content (between residues 770–1110) and seems not to contain a stably folded domain according to limited proteolysis experiments (Fath et al., 2007). Consistent with this, the majority of Sec31 fragments that we prepared were heavily degraded during expression in E. coli as fusions to glutathione S-transferase. We were able to express and purify a fragment encompassing most of the proline-rich region, residues 879–1114 (construct A in Figure 2B). Figure 2A shows the results of the experiment comparing the activities of this fragment and full-length Sec13/31 across a range of concentrations. The maximal GTPase rate for both proteins was 0.1 s⁻¹; the apparent Kₘ for the proline-rich region binding to Sec23/24•Sar1 was 6.0 μM, just slightly weaker than the value of 4.8 μM measured for full-length Sec13/31. Thus, the proline-rich region retains the binding and catalytic capacity of full-length Sec13/31. This effect is exclusive to the proline-rich region, as other domains of Sec31 had no effect on GAP activity (data not shown).

Next, we tested a series of Sec31 truncation constructs for the ability to stimulate GAP activity (Figures 2B and 2C). All the Sec31 protein fragments were added at 1.5 μM to an assay mix containing 1 μM Sar1-mant-GTP and 3 μM Sec23/24. Under these conditions, full stimulatory activity by Sec31 yielded an ~5-fold rate enhancement relative to the absence of Sec31 (compare, for example, the rate due to construct A with the control rate in Figure 2C). The results of the experiment show that an ~50 residue polypeptide, residues 899–947, retains full activity (construct K in Figures 2B and 2C). We refer to this as the active fragment of Sec31. Removal of additional residues from the N or C termini causes a reduction in activity (fragments D, E, and J). However, there is not the dramatic loss of activity that one would predict if the terminal residues of a folded domain were removed. Consistent with this, the crystallographic analysis reveals that the Sec31 active fragment does not fold into a compact structure to interact with Sec23•Sar1, but binds as an extended polypeptide across the surface of the inner-shell complex (see below).

Surprisingly, when we tested the Sec31 active fragment across a range of concentrations, we found that it was more effective than full-length Sec13/31 at stimulating the GAP activity of Sec23/24. As shown in Figure 2D, the maximal rate of GTP hydrolysis measured with the active fragment was three times faster than for full-length Sec13/31 (0.3 s⁻¹ compared to 0.1 s⁻¹). Yet the active fragment lost a corresponding 3-fold affinity for Sec23/24•Sar1 relative to the full-length protein. This subtle difference in stimulatory activity is perhaps not surprising given the very different forms of the two molecules—a 5 kDa peptide versus an ~340 kDa heterotetramer. Regardless, the data establish that GAP stimulatory activity is due to a circumscribed peptide element located in the middle of the Sec31 proline-rich region (Figure 2B).

Finally, the most rapid GTPase rate that we observed in this study—0.24 s⁻¹ measured using 40 μM active fragment—was thirty times faster than the rate due to Sec23/24 alone (Figure 2D). This ~30-fold stimulation of Sec23/24 GAP activity by the Sec31 active fragment compares to the 10-fold stimulation reported by Antonny et al. (2001) using liposomes and membrane-bound coat proteins.

### Sec31 Does Not Contact Sec24 to Accelerate GTP Hydrolysis

The small size of the Sec31 active fragment suggests that it does not contact all three subunits of the Sec23/24•Sar1 complex. On the other hand, two-hybrid analysis identified two interactions involving the Sec31 proline-rich region—one with Sec23 and the other with a fragment of
Sec24 (Shaywitz et al., 1997). We tested the contribution of Sec24 to the Sec31-stimulated GAP reaction (Figure 2D). First, we used the Sec31 active fragment, and found that Sec23 was almost as effective as Sec23/24 at catalyzing GTP hydrolysis on Sar1 (crosses and triangles in Figure 2D). There is a very slight loss of affinity of Sec31 for the inner shell proteins that may reflect experimental error (apparent \( K_m \) for Sec31 binding to Sec23/24+Sar1 is 4.8 \( \mu M \)). For the Sec31 proline-rich region, the maximal GTPase rate is also 0.10 s\(^{-1}\), and the apparent \( K_m \) for the 236 residue region binding to Sec23/24+Sar1 is 6.0 \( \mu M \).

Next, we tested full-length Sec13/31 and found that Sec23 was as effective as Sec23/24 in the GAP reaction. Sec24 had no effect on the catalytic rate or the affinity of Sec13/31 for the inner-shell proteins (circles and diamonds in Figure 2D). These results suggest that Sec24 is not involved in the contacts between the inner and outer shells. It is conceivable that an interaction between Sec13/31 and Sec24 occurs but is somehow not reported by our solution-based GAPase assay, though we consider this unlikely. The evidence for an interaction between Sec31 and Sec24 was obtained from a two-hybrid experiment employing a C-terminal region of Sec24, residues 666–926 (Shaywitz et al., 1997). According to the Sec24 crystal structure (Bi et al., 2002), the 666–926 fragment will have portions of the hydrophobic core of its helical domain exposed to solvent and may not be a suitable interactor for two-hybrid analysis.
In summary, these data define a COPII core complex that stimulates rapid GTP hydrolysis, comprising Sec23•Sar1 and the active fragment of Sec31. We assembled this complex from purified S. cerevisiae proteins and obtained crystals that diffract X-rays to 2.5 Å resolution (Table 1).

<table>
<thead>
<tr>
<th>Table 1. Data Collection and Refinement Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space group</strong></td>
</tr>
<tr>
<td><strong>Cell dimensions (Å)</strong></td>
</tr>
<tr>
<td><strong>PDB accession #</strong></td>
</tr>
<tr>
<td><strong>Data processing</strong></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
</tr>
<tr>
<td><strong>Measured reflections</strong></td>
</tr>
<tr>
<td><strong>Unique reflections</strong></td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
</tr>
<tr>
<td><strong>I/σ</strong></td>
</tr>
<tr>
<td><strong>R_merge (%)</strong></td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
</tr>
<tr>
<td><strong>Data range (Å)</strong></td>
</tr>
<tr>
<td><strong>Reflections</strong></td>
</tr>
<tr>
<td><strong>Nonhydrogen atoms</strong></td>
</tr>
<tr>
<td><strong>Water molecules</strong></td>
</tr>
<tr>
<td><strong>R.m.s. Δ bonds (Å)</strong></td>
</tr>
<tr>
<td><strong>R.m.s. Δ angles (%)</strong></td>
</tr>
<tr>
<td><strong>Mean B factor (all atoms)</strong></td>
</tr>
<tr>
<td><strong>Mean B factor (Sec31 atoms)</strong></td>
</tr>
<tr>
<td><strong>R factor (%)</strong></td>
</tr>
<tr>
<td><strong>R_free (%)</strong></td>
</tr>
</tbody>
</table>

*Rmerge = 100 x ∑_i ∑_j | hi,j - <hi> | / ∑_i ∑_j hi,j.*

**DISCUSSION**

**Architecture of Sec23•Sar1 Bound to the Active Fragment of Sec31**

The overall conformations of Sec23 and Sar1 observed in the crystal structure of the ternary complex (Figures 3 and 4) are essentially the same as in the Sec23•Sar1 binary complex (Bi et al., 2002), with only subtle alterations observed in the vicinity of the Sar1 active site. Sar1 is stabilized in the GTP conformation by the nonhydrolyzable analog, such that its β2-β3 switching element has shifted to eliminate the binding site for the Sar1 N-terminal “membrane anchor” (Bi et al., 2002).

Thirty-six residues of Sec31, residues 907–942, are included in the molecular model, as no electron density is observed for the eight N-terminal and five C-terminal positions. This is consistent with the results of the biochemical dissection (see Figures 2B and 2C), and suggests that as few as 35–40 residues of Sec31 are required for GAP stimulation.

The Sec31 active fragment binds as an extended poly-peptide across a composite surface of the Sec23 and Sar1 molecules. It does not adopt a folded tertiary structure; moreover, it has hardly any secondary structure other than a single turn of α helix, residues 915–919, that binds to the switch 2 element of Sar1 (Figure 3). The active fragment is oriented with its N terminus bound to Sar1 and C terminus bound to Sec23. A 16 residue segment (907–922) that includes the turn of α helix, extends ~15 Å across the Sar1 surface, interacting with switch 2 and the adjacent α3 helix (a common binding site for GAPs on G proteins [Vetter and Wittinghofer, 2001]). The 20 C-terminal residues 923–942 bind in a highly extended conformation, 47 Å long, across Sec23, and contact the gelsolin, trunk, and β-barrel domains of the Sec23 molecule (Bi et al., 2002).

The observation that the Sec31 active fragment binds to a composite surface of Sec23 and Sar1 provides a molecular explanation for the ordered recruitment of the Sec23/24 and Sec13/31 complexes to ER membranes. Only upon GTP-dependent formation of Sec23/24•Sar1 is the receptor site for Sec13/31 formed on the membrane (Barlowe et al., 1994; Matsuoka et al., 1998). A structural corollary of this is the layered arrangement of the membrane and proteins in the coat, as highlighted in Figure 4. Thus, the Sec31 active fragment binds to the membrane-distal surface of Sec23•Sar1. All of its residues reside a uniform distance from the membrane surface, which we estimate to be ~45 Å (this assumes that the membrane-proximal surface of Sec23/24•Sar1 closely apposes the phospholipid membrane).

The sequence composition of the active fragment (Figure 5A) strongly suggests that the isolated peptide will adopt an unstructured conformation in solution. Only residue Trp922 of the ~40 amino acids is conserved as a large

---

**Crystal Structure Determination of Sec23•Sar1 Bound to the Active Fragment of Sec31**

The crystals of the ternary complex of Sec23•Sar1 and the Sec31 active fragment were grown using the truncated form of Sar1 (residues 24–190) and were stabilized using the nonhydrolyzable GTP analog GppNHp. The same strategy was used previously to crystallize the Sec23•Sar1 complex (Bi et al., 2002). Thus, the binary and ternary crystal structures can be compared straightforwardly to assess structural differences induced by Sec31.

The crystal structure of the ternary complex was solved by the molecular replacement method using yeast Sec23 and Sar1 as search models. The Sec31 active fragment was built into strong electron density for residues 907–942 of the peptide, and the structure was refined using data to 2.5 Å resolution (Figure 3 and Table 1).
hydrophobic side chain. Moreover, a limited proteolysis analysis of Sec13/31 did not identify a stably-folded domain in this region or in the proline-rich region of Sec31 (residues 770–1110) as a whole. Whether the Sec31 active fragment is thoroughly unstructured in the context of full-length Sec13/31 is unclear. The 3-fold differences in catalytic rate and affinity on Sec23/24/C15 Sar1 that we reported above may hint at regulatory conformational constraints imposed on the active fragment by adjacent domains of Sec13/31, but the development of this speculative idea must await further experimental work.

Interfacial Contacts

In Figure 3, the Sec31 active fragment is colored to delineate five segments. The N-terminal residues, 907–919 (colored purple), interact solely with Sar1. These residues are not conserved in Sec31 sequences (Figure 5A), but we observe close interactions between this region and the switch 2 and α3 elements of Sar1. Next, residue Asp920 and the highly conserved residues Gly921 and Trp922 of the active fragment (colored white in Figure 3) contact both Sec23 and Sar1 to form a tripartite protein interface near the Sar1 active site (Figure 5D). The peptide geometry around residue Gly921 guides the insertion of the Trp922 side chain close to the seat of reaction (Figure 5D), and these interactions are key to the stimulation of GAP activity, as described in more detail below.

The remainder of the active fragment, residues 923–942, interacts solely with Sec23. A central stretch, residues 928–934 (yellow in Figure 3), appears to interact loosely, whereas the regions on either side form more intimate interactions with Sec23. Residues 923–927 (blue in Figure 3) are well conserved among Sec31 sequences (Figure 5A), and form a series of interactions with Sec23 that appear to be important for affinity and for buttressing Trp922 at the Sar1 active site (Figure 5B and D). Finally, the C-terminal stretch comprising residues 935–942 (green in Figure 3) is not well conserved, but several side chains—in particular Ala936, Val939, and Val941—form intimate contacts with residues of the trunk and β-barrel domains of Sec23 (Figure 5B).

In summary, the binding site for the Sec31 active fragment extends ~60 Å across the surface of Sec23 and Sar1, and
involves three quasi-independent binding regions (colored purple, white/blue, and green in Figure 3). Most residue positions in the Sec31 active fragment are not well conserved, in particular the N-terminal sequence that interacts with Sar1 and the C-terminal portion that interacts with Sec23. Nevertheless, these two terminal interaction regions are important for binding and stimulation of GAP activity, according to the results of the dissection experiment (Figure 2C). Sequence conservation is restricted to a central set of residues of the active fragment that is clustered around the tripartite interface near to the Sar1 active site.

Sec31 Residues Complete the Sar1 Active Site for Rapid GTP Hydrolysis

Sar1 has a very slow intrinsic rate of GTP hydrolysis because, like other Ras proteins, it lacks key catalytic residues (Bi et al., 2002; Vetter and Wittinghofer, 2001). The mechanism by which Sec32 acts as a GAP to accelerate the reaction involves the insertion of an arginine side chain, Arg722, into the active site to form bonds to the phosphates via its guanidinum group (Bi et al., 2002; Figure 5C). This type of mechanism, involving an “arginine finger” residue that neutralizes negative charge in the GTPase transition state is likewise a common feature of Ras proteins (Vetter and Wittinghofer, 2001).

The crystal structure of the Sec32●Sar1 binary complex revealed two catalytic features in addition to the identity of the arginine finger residue. First, an extensive interface between Sec32 and the switch 1 and 2 elements of Sar1 stabilizes these regions close to the active site. Second, a water molecule bridges the imidazole side chain of His77 and the \( \gamma \)-phosphate group of GppNHp, and is suitably located for nucleophilic attack (Bi et al., 2002; Figure 5C). Taken together, the catalytic features of the Sec32●Sar1 active site are very similar to those seen in other GAP●G protein complexes, so it is not immediately obvious from the Sec32●Sar1 crystal structure how Sec31 might stimulate GAP activity (discussed in Bi et al., 2002).

Inspection of the Sar1 active site and the tripartite interface in the ternary complex now reveals how Sec31 stimulates GAP activity of Sec32 (Figure 5D). The active fragment inserts two residues, Trp922 and Asn923, close to the active site, with the plane of the indole ring of Trp922, into the active site to form bonds to the phosphates via its guanidinum group (Bi et al., 2002; Figure 5C). This type of mechanism, involving an “arginine finger” residue that neutralizes negative charge in the GTPase transition state is likewise a common feature of Ras proteins (Vetter and Wittinghofer, 2001).
is -60° in the binary complex and -75° in the ternary complex). Other changes at the active site induced by Sec31 binding appear to support the role of Trp922. In particular, Sec23 residue Gln720, which is oriented away from the active site in the binary complex, is turned toward the active site and forms hydrogen bonds to residues Trp922 and Asn923 of the active fragment. Thus, Sec31 side chains do not have a chemical catalytic role at the Sar1 active site like the arginine finger residue of Sec23—indeed, only residue Pro926 is invariant.

Figure 5. Sec31 Interactions at the Sec23•Sar1 Interface and the GTPase Active Site
(A) Alignment of eight sequences of the Sec31 active fragment from seven species (including human forms A and B). The key at the top indicates the conservation of the predominant residue; two black bars correspond to two common occurrences out of eight—the “noise level”—and red bars highlight the more highly conserved positions. Key tryptophan and asparaginase residues are indicated with stars.
(B) Schematic drawing showing contacts at the protein-protein interfaces, colored as in Figures 3 and 4. Select contacts that are of structural interest are indicated with black lines. The contact between the arginine finger residue of Sec23—Arg722 (labeled R722)—and GTP phosphate groups is indicated with an arrow. The negative charge of the side chain of Sec31 residue D924 interacts with the electrostatic dipole of helix αI on Sec23. Residue Phe380 on helix αI is mutated to leucine in human Sec23A in individuals with cranio-lenticulo-sutural dysplasia (Boyadjiev et al., 2006).
(C) Bar graph shows the results of the mutagenesis experiment. The Sec31 fragment (residues 899–947) and mutants thereof were tested for the ability to stimulate the GAP activity of Sec23/24. In all experiments the Sec31 polypeptide was present at 1.5 μM, Sar1-mantGTP at 1 μM, and Sec23/24 at 3 μM. (These conditions are the same as in the dissection experiment shown in Figure 2C). GAP stimulatory activity is expressed as a percentage of the activity of the active fragment. The standard errors were obtained from least squares fits to the data.
(D) Close-up view showing Sec23 contacts at the Sar1 active site. This picture is generated from a previously determined crystal structure of Sec23 bound to Sar1 and GppNHp. Note the orientation of the arginine finger residue R722 and its bonds to the nucleotide.
(E) Close-up view of the ternary complex is oriented as in (C). The difference electron density map was calculated prior to the inclusion of the Sec31 active fragment (residues 899–947) and mutants thereof were tested for the ability to stimulate the GAP activity of Sec23/24. In all experiments the Sec31 polypeptide was present at 1.5 μM, Sar1-mantGTP at 1 μM and Sec23/24 at 3 μM. (These conditions are the same as in the dissection experiment shown in Figure 2C). GAP stimulatory activity is expressed as a percentage of the activity of the active fragment. The standard errors were obtained from least squares fits to the data.
in the subset of sequences shown in Figure 5A. Rather, GAP stimulation is likely caused by Trp922 of Sec31 interacting with His77 to orient its imidazole ring at the Sar1 active site. In so doing, Trp922 and Asn923 plug a solvent-filled cavity that extends from bulk solvent to the vicinity of the active site, an unfavorable arrangement for catalysis in the binary complex that leaves one surface of the indole ring of His77 exposed to solvent (Figure 5C).

Structure-based mutagenesis of the Sec31 peptide confirmed the importance of these key residues. Mutation of residue Trp922 and Asn923 to alanine caused complete loss of GAP stimulatory activity (Figure 5C). Likewise, changes to Leu925 and residue Val939 in the C-terminal region were highly disruptive, even though these are not highly conserved positions (Figure 5A). Two very conservative residue changes in the N-terminal portion of the active fragment, Q910A and N915A, were tolerated with only modest loss of stimulatory activity. The asparagine residue does not in fact contact Sec23 or Sar1 directly.

In summary, the active-site configuration of the ternary complex explains how Sec31 synergizes with Sec23 to accelerate GTP hydrolysis. Sec31 cannot act alone on Sar1 because it only binds to the Sec23•Sar1 complex. In this way, the sequential binding reaction confers a two-gear mechanism for GTP hydrolysis on Sec13 (Antony et al., 2001), whereby hydrolysis is initiated upon Sec23/24 binding and is accelerated further upon recruitment of Sec13/31. Thus, GTP hydrolysis is programmed into the COPII system upon assembly. The slow rate of reaction on Sec23/24•Sar1 may provide the prebudding complex the opportunity to gather cargo and SNAP receptors and move the cargo on the membrane towards the vesicle. Since each assembly unit is a Sec13/Sec31 heterotrimer, there are 48 copies of Sec31 per cage, and binding sites for 48 copies of Sec23/24•Sar1. Put differently, the cage has twelve vertices, each vertex formed from the convergence of four Sec13/31 rods, so there are binding sites for four Sec23/24•Sar1 molecules to nestle on the membrane underneath each vertex.

With respect to the present study, there are three salient features of the model. First, the membrane vesicle is drawn as a 40 nm sphere inside the cage, to allow an approximate fashion with the Sec31–Sar1 complex explains how the COPII coat is probably restricted to the inner shell proteins. Nevertheless, the mobility of Sec23/24•Sar1 is probably restricted by its dense packing on the membrane surface. We previously estimated that 48 copies of Sec23/24•Sar1 would cover as much as 80% of the surface area of a 40 nm membrane vesicle (Fath et al., 2007).

We have depicted this situation as four Sec23/24•Sar1 complexes under a vertex but not conforming closely to the cage symmetry (Figure 6). The complexes are drawn nested close to the vertex center, and paired in an approximate fashion with the Sec31 α-solendoid domain to

The F382L Disease Mutation in Human Sec23A Maps Close to the Interface of Sec23 and Sec31

A substitution in human Sec23A, F382L, causes cranio-lenticulo-sutural dysplasia, a craniofacial and skeletal dysmorphic syndrome (Boyadjiev et al., 2006). Fibroblasts homozygous for the mutation have a disorganized endoplasmic reticulum. Molecular analysis of F382L Sec23A reveals that although it can combine with Sec24 and is recruited efficiently to membranes by Sar1, the mutant protein is deficient in vesicle formation. Moreover, Sec13/31 is mislocalized to the cytoplasm in the mutant cells, suggesting that the mutation affects the interaction of Sec23/24•Sar1 with Sec13/31 (Boyadjiev et al., 2006, and see accompanying manuscript by Fromme et al., 2007).

The corresponding phenylalanine residue in S. cerevisiae Sec23, Phe380, is located on helix α1 of the trunk domain, which makes key interactions with the Sec31 active fragment (Figure 5). The phenylalanine residue is highly conserved (Boyadjiev et al., 2006), and its side chain resides on the internal face of helix α1, contributing to the hydrophobic core via contacts to adjacent hydrophobic side chains of Sec23, including Phe346, Phe375, and Tyr384. Although Phe380 does not contact Sec31 directly, residues 924–931 of the active fragment make contacts along the length of helix α1, the most notable of which is the electrostatic interaction between the side chain of Asp924 and the positive charge of the N terminus of the α helix (Figure 5D). As noted above, this set of residues of the active fragment, in particular 923–927, are important for buttressing Trp922 at the Sar1 active site. The disease mutation is a subtle change to leucine, but we predict that this will perturb the local structure of helix α1 so as to weaken contacts to the active fragment. We infer that the Sec23 disease mutation impairs the recruitment and nucleation of Sec13/31 at sites of COPII budding.

Connection of Inner and Outer Shell Complexes in the COPII Coat

Figure 6 presents a model for the arrangement of Sec23/24 complexes in the COPII coat, and is based on recent structural data on the 60 nm cuboctahedron cage built from 24 assembly units (Fath et al., 2007; Stagg et al., 2006). Since each assembly unit is a Sec13/Sec31•Sec13/Sec13 heterotrimer, there are 48 copies of Sec31 per cage, and binding sites for 48 copies of Sec23/24•Sar1. Put differently, the cage has twelve vertices, each vertex formed from the convergence of four Sec13/31 rods, so there are binding sites for four Sec23/24•Sar1 molecules to nestle on the membrane underneath each vertex.

With respect to the present study, there are three salient features of the model. First, the membrane vesicle is drawn as a 40 nm sphere inside the cage, to allow an approximate fashion with the Sec31–Sar1 complex explains how the COPII coat is probably restricted to the inner shell proteins. Nevertheless, the mobility of Sec23/24•Sar1 is probably restricted by its dense packing on the membrane surface. We previously estimated that 48 copies of Sec23/24•Sar1 would cover as much as 80% of the surface area of a 40 nm membrane vesicle (Fath et al., 2007).

The flexible connection between the cage and Sec23/24•Sar1, via the 130 residue linker, suggests that Sec23/24•Sar1 complexes may be somewhat mobile on the membrane, such that the 432 symmetry of the cage is not imposed strictly on the inner shell proteins. Nevertheless, the mobility of Sec23/24•Sar1 is probably restricted by its dense packing on the membrane surface. We previously estimated that 48 copies of Sec23/24•Sar1 would cover as much as 80% of the surface area of a 40 nm membrane vesicle (Fath et al., 2007).

We have depicted this situation as four Sec23/24•Sar1 complexes under a vertex but not conforming closely to the cage symmetry (Figure 6). The complexes are drawn nested close to the vertex center, and paired in an approximate fashion with the Sec31 α-solendoid domain to
Figure 6. A Model for the Interactions and Organization of Proteins in the COPII Coat
The picture shows a model for the COPII cage, a cuboctahedron built from 24 copies of the Sec13/Sec31-Sec31/Sec13 assembly unit (Fath et al., 2007; Stagg et al., 2006). Each of the twelve vertices of the cage is formed by the convergence of four Sec13/31 subunits. Thus, there are binding sites for four copies of Sec23/24/Sar1 under each vertex. In the picture we have modeled four copies of Sec23/24/Sar1 complexed with the Sec31 active fragment. The zigzag lines represent the sites for four copies of Sec23/24/Sar1 as it does from the arrangement and symmetry of the outer shell assembly units.

A flexible connection between the inner and outer shells of the vesicular coat is highly reminiscent of the situation in clathrin cages, where the N-terminal b-propeller domain of the clathrin heavy chain interacts with adaptors via short peptide elements that are nested in flexible regions of polypeptide (ter Haar et al., 2000). Thus, a flexible connection between the architectural outer shell and the cargo-gathering inner shell complex may be a common feature of vesicle coat organization that facilitates the packaging of a range of cargo molecules of different shapes and sizes.

EXPERIMENTAL PROCEDURES
Protein Preparation
For protein expression in insect cells, baculoviruses (Bac-to-Bac, GIBCO) were prepared encoding yeast Sec13 and His6-Sec31. These proteins, as well as yeast His6-Sec23 and Sec23/His6-Sec24 complex, were expressed in insect cells and purified as described (Bi et al., 2002). Histidine tags were removed by cleavage with TEV protease. The soluble, truncated form of yeast Sar1, comprising residues 24–190, was complexed with the nonhydrolyzable GTP analog guanosine-5′-[β,γ-imido]-triphosphate (GppNHp) and Mg2+, and purified as described (Bi et al., 2002). Fragments of Sec31 were subcloned according to the domain analysis of Fath et al., 2007, and were over-expressed in E. coli as C-terminal fusions to glutathione S-transferase using the pETGEXCT vector (Sharovcks, 1994). Proteins were purified by glutathione sepharose chromatography. Mutagenesis was performed using the QuikChange reagents (Stratagene).

Crystallization and Structure Determination
For crystallization, Sec23 was mixed with Sar1•GppNHp•Mg2+ (truncated Sar1) and the active fragment of Sec31 (residues 899–947). Each protein was at a final concentration of 360 μM in 150 mM NaCl, 20 mM HEPES (pH 7.4), and 4 mM DTT. Crystals (space group P21212) were grown at 22°C by the hanging-drop method by mixing 1 μl of protein solution with 1 μl of a well solution comprising 15% PEG-1500 and 100 mM HEPES (pH 7.0). Crystals, which appeared after three days, were transferred to a solution containing 20% PEG-1500, 16% glycerol, and 100 mM HEPES (pH 6.8), and flash-frozen in liquid propane. Crystals treated in this manner diffracted X-rays to at least 2.5 Å resolution using a synchrotron X-ray source.

X-ray diffraction data to 2.5 Å resolution were measured at beamline X-25 of the National Synchrotron Light Source (NSLS). Data were processed with programs DENZO and SCALEPACK (Otwinowski and Minor, 1997). The structures were solved by molecular replacement with the program AMORE (CCP4, 1994) using yeast Sec23 as the search model (Bi et al., 2002). The initial model was improved by rigid-body and positional refinement with program CNS (Brünger et al., 1998). Sar1•GppNHp•Mg2+ was then placed into clear electron density. Finally, residues 907–942 of Sec31 were modeled into electron density, and the structure was refined to an R factor of 20.5% (Rfree = 26.2%). The X-ray data and refinement statistics are summarized in Table 1. The model contains three residues of Sec23 that are outliers in the Ramachandran plot: Phe659 seems to be modeled incorrectly; residues Val30 and Gln683 are located in regions of weak electron density. None of these residues is in the vicinity of the interatomic contacts with Sar1 or Sec31. All Sec31 residues lie within allowed regions of the Ramachandran plot. The final model comprises 7354 protein atoms, one molecule of GppNHp•Mg2+, one zinc atom (in the zinc finger domain of Sec23) and 231 water molecules. No electron density is observed for the eight N-terminal and five C-terminal residues of the Sec31 polypeptide.

GTPase Assays
For GTPase assays, we prepared the substrate Sar1 (24–190) bound to the fluorescent nucleotide derivatives mant-GTP, mant-dGTP or mant-GppNHp (mant nucleotides were purchased from Jena Bioscience).
Sar1 was incubated with a 30-fold molar excess of mant nucleotide overnight at 4°C, and excess nucleotide was separated from protein by gel filtration on a Superdex 75 column. The extent of mant nucleotide incorporation was estimated by absorbance measurements at 280 and 355 nm.

The decrease in fluorescence, due to hydrolysis of mant-GTP on Sar1, was monitored in a fluorimeter (Fluoromax-2, Horiba Jobin-Yvon). All reactions were performed at 25°C in 150 mM NaCl, 25 mM HEPES (pH 7.4), 2 mM DTT, and 2 mM MgCl₂ using an excitation wavelength of 360 nm and emission measured at 438 nm. In a typical experiment, the reaction was initiated by the addition of Sec23 (or Sec23/24) to a filtered, degassed solution containing Sar1•mant-nucleotide plus Sec31; the total reaction volume was 1.2 ml. Reaction rates were determined from exponential fits to the data.

ACKNOWLEDGMENTS
We thank Anand Saxena for assistance at synchrotron beamline X25 at NSLS. This work was supported by grants from the National Institutes of Health and the Howard Hughes Medical Institute.

REFERENCES

Accession Numbers
The coordinates have been deposited in the Protein Data Bank with the accession code 2QTV.