Structure of the Sec23/24–Sar1 pre-budding complex of the COPII vesicle coat

Xiping Bi, Richard A. Corpina & Jonathan Goldberg

Howard Hughes Medical Institute and the Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

COPII-coated vesicles form on the endoplasmic reticulum by the stepwise recruitment of three cytosolic components: Sar1–GTP to initiate coat formation, Sec23/24 heterodimer to select SNARE and cargo molecules, and Sec13/31 to induce coat polymerization and membrane deformation. Crystallographic analysis of the *Saccharomyces cerevisiae* Sec23/24–Sar1 complex reveals a bow-tie-shaped structure, 15 nm long, with a membrane-proximal surface that is concave and positively charged to conform to the size and acidic-phospholipid composition of the COPII vesicle. Sec23 and Sar1 form a continuous surface stabilized by a non-hydrolysable GTP analogue, and Sar1 has rearranged from the GDP conformation to expose amino-terminal residues that will probably embed in the bilayer. The GTPase-activating protein (GAP) activity of Sec23 involves an arginine side chain inserted into the Sar1 active site. These observations establish the structural basis for GTP-dependent recruitment of a vesicular coat complex, and for uncoating through coat-controlled GTP hydrolysis.

Recent research on vesicle transport pathways responsible for the growth and maintenance of organelles has shown that vesicle budding and fusion mechanisms are conserved in organisms from yeast to humans[1]. Budding occurs by the action of cytoplasmic coat protein complexes (COPs) that polymerize on the membrane surface, capturing cargo and SNARE molecules in the process, and deforming the membrane to sculpt out vesicles[2]. Small GTP-binding proteins of the ARF family regulate coat assembly and disassembly by providing both the spatial cue—activated ARF–GTP—that restricts coat recruitment to specific export sites[3–5], and the temporal cue—controlled GTP hydrolysis—that triggers uncoating[6–8]. Cells contain a variety of COPs, including COPI or coatomer, COPII, and numerous isoforms of clathrin/adaptin complexes[2], each COP specifying a discrete transport step according to its affinities for a particular export site and for a particular set of SNARE molecules gathered up during budding to programme the vesicle for fusion with its target membrane[9–11].

COPII-coated vesicles bud from the endoplasmic reticulum (ER) to export newly synthesized proteins to the Golgi complex. The components of COPII comprise the ARF-related molecule Sar1 and two large heterodimeric complexes, Sec23/24 and Sec13/31 (ref. 5). These five proteins when supplied with GTP can bud characteristic 50–90-nm vesicles from synthetic liposomes, provided these contain about 10% acidic phospholipids for effective binding of Sec23/24 (ref. 12). *In vivo*, budding is initiated by the exchange of GDP for GTP on Sar1 catalysed by Sec12, an ER-localized guanine-nucleotide exchange factor, or GEF[13–15]. Sar1–GTP then binds to the membrane, probably by embedding an N-terminal α-helix[16] in the bilayer in the manner delineated for ARF1 (refs 17–20). Sar1–GTP recruits Sec23/24 through an interaction with the Sec23 subunit[7], and this pre-budding complex binds Sec13/31, which is believed to act, like clathrin, to polymerize the coat and deform the membrane into a bud[5,12]. Electron microscopy (EM) of isolated complexes shows that Sec23/24 resembles a bow tie and that Sec13/31 is elongated and apparently rather flexible[21]. EM analysis of the COPII vesicle coat suggests a layered structure comprising an inner shell of Sec23/24 and an outer shell of Sec13/31 (ref. 22), recapitulating the sequential binding mechanism observed in budding experiments.

Studies of the cargo-selection function of Sec23/24–Sar1 in yeast and mammalian systems have shown that ER-to-Golgi SNAREs and at least some integral membrane cargo proteins can bind directly to the pre-budding complex[10,23–26], leading to their concentration in COPII vesicles[27,28]. A final function of Sec23/24—one that highlights the dynamic nature of vesiculation—is as the GAP for the otherwise inert GTPase activity of Sar1–GTP[7]. It seems that, as the pre-budding complex forms, a slow GTP hydrolysis reaction is programmed into the assembling coat to limit its lifetime and to couple coating to uncoating[7,8].

In the present study, molecular aspects of COPII function in ER export were addressed by determining the structure of the Sec23/24–Sar1 complex. Its domain composition and architecture bear no resemblance to AP2 adaptin or clathrin[29,30]. An extensive membrane-interaction area is formed by all three proteins and is concave and basic to match the COPII vesicle surface. The nature of the interface between Sec23 and Sar1 explains how GTP binding initiates coat recruitment and how coat recruitment initiates GTPase activity, suggesting a molecular-level picture of the control of coat assembly and disassembly by a cycle of GTP binding and hydrolysis.

Structure determination

All structural studies used full-length recombinant yeast Sec23 and Sec24 proteins. Because full-length Sar1–GTP tends to aggregate in solution[31], a truncated form was used that lacks the N-terminal 23 amino acids. Biochemical studies show that ARF proteins truncated in this way specifically lose the membrane-interaction function and retain all protein-interaction functions[32,33]; structural studies of ARF proteins in truncated and full-length forms show that N-terminal truncation has no evident effect on conformation[19,34,35].

The atomic model of Sec23/24–Sar1 was built from structures of Sec23/24 and Sec23–Sar1 determined by X-ray crystallography. Crystals of the yeast Sec23–Sar1 complex (relative molecular mass (M_r) of Sec23 85K, Sar1 19K) bound to the non-hydrolysable GTP analogue 5′-guanylyl imidodiphosphate (GppNHp) and Mg^{2+} grew in the space group $P2_12_12_1$ with two complexes in the asymmetric unit. The initial electron density map was calculated to 2.7 Å resolution from a multiwavelength anomalous diffraction

(MAD) data set incorporating selenium as the anomalous scatterer. The structure was refined with native X-ray data to 2.5 Å resolution (Supplementary Information Table 1). The yeast Sec23/24 complex (Sec24 M_r 104K) crystallized in the space group $P2_12_12_1$. This structure was determined by the molecular replacement method, and refined with data to 2.75 Å resolution (Supplementary Information Table 1).

Comparison of the two crystal structures shows that only minor changes occur in Sec23 upon interaction with either Sar1 or Sec24. In addition, the binding sites on Sec23 for the two proteins are separated by more than 25 Å, and hence it was possible to construct a straightforward composite model of Sec23/24–Sar1 from the two experimental structures.

Architecture and membrane interaction

Figure 1a shows the inner face of the coat that will probably appose the membrane, and highlights the bow-tie shape of the complex. The folds of Sec23 and Sec24 are closely related, and the two subunits interact about a dyad (Fig. 1a is viewed along this axis). Comparison with the 'side' view (Fig. 1b) shows that Sec23/24–Sar1 is elongated in directions parallel to the membrane so as to cover an extensive membrane area; it measures ~150 Å in the long direction, and varies along the bow tie from ~75 Å at the Sec24 end, to as little as 30 Å at the dimer interface, to ~100 Å at the Sec23–Sar1 end (Fig. 1a). By contrast, the complex is relatively thin and extends uniformly only about 40 Å in the direction away from the membrane (Fig. 1b), except for two protrusions formed by a gelsolin-like domain on each subunit that extend ~60 Å (Figs 2a and 3a). For comparison, the thickness of the COPII coat is ~100 Å, on the basis of EM images[22].

A distinctive feature of the pre-budding complex is its concave inner surface, which seems to match the size of a 60-nm vesicle (Fig. 1b and central images in Fig. 3). Although Sec13/31 binding is thought to provide the driving force for vesiculation, the observed curvature suggests that Sec23/24–Sar1 will favour membrane deformation and perhaps exert some influence on the final dimensions and rather uniform size of COPII vesicles that bud from purified ER membranes[5]. Consistent with this idea, *in vitro* experiments show that the yeast Sec24 homologue Lst1 can combine with Sec23 to form mixed Sec23/24–Sec23/Lst1 coats that bud somewhat larger (~12 nm greater in diameter) and size-heterogeneous vesicles[36].

Sec23/24–Sar1 appears complementary to the vesicle surface both in shape and chemistry. Examination of the protein electrostatic surface reveals that the concave, inner face is generally positively charged, and the outer face much more acidic (Fig. 3a). This polarization is conserved from yeast to human Sec23/24 complexes (as a general conservation of the surface regions rather than invariance of numerous arginine and lysine side chains). These observations can explain the results of *in vitro* budding experiments that demonstrated a requirement for acidic phospholipids for the Sar1-dependent binding of Sec23/24 to synthetic liposomes[12].

The Sec23/24–Sar1 structure reveals three determinants of membrane interaction: the curved form and positively charged surface as described above, plus the orientation of the Sar1 N terminus within the complex. Sar1–GppNHp interacts extensively with Sec23 to form a continuous surface that extends the membrane-interaction area (Figs 1a and 2a), and this surface includes the Sar1 N terminus (Gly24 is the first modelled residue), which projects towards the membrane, probably to embed N-terminal residues in the bilayer (see below). If Sec23/24–Sar1 interacts closely with the membrane as these observations suggest, then the concave inner surface of one complex would cover ~8,200 Å^2 of membrane, roughly 0.7% of the

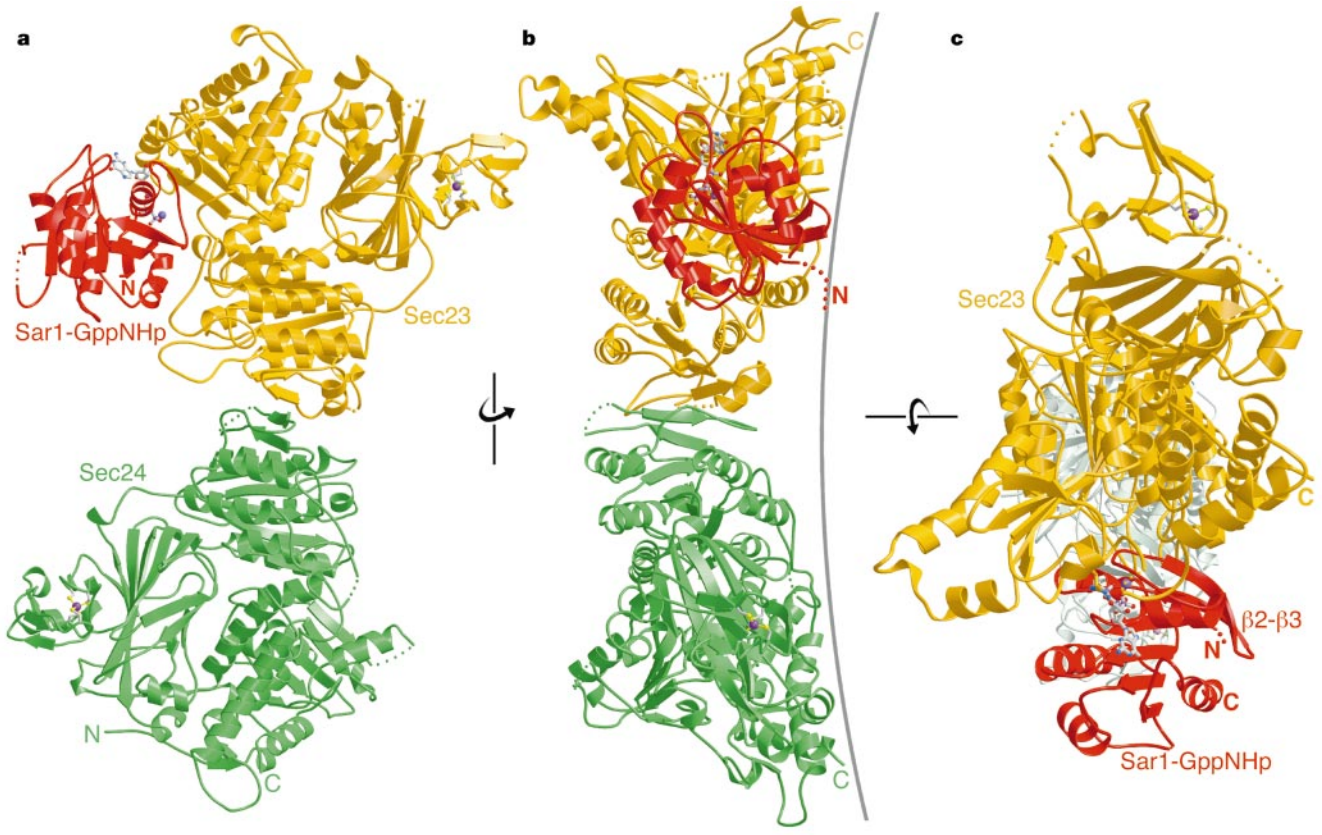

Figure 1 Structure of the Sec23/24–Sar1 pre-budding complex. Ribbon representation is shown as successive 90° rotations. Sec23 is yellow, Sec24 green and Sar1 red. **a**, Front view along the dyad relating the coat subunits, with the membrane-proximal surface facing forward. **b**, Side view. The grey line indicates the curvature of a 60-nm COPII vesicle, drawn to scale. **c**, Top view. The three regions of Sar1 that would face the membrane are labelled C, N and β2–β3.

surface area of a 60-nm vesicle; and 60 such complexes would provide ~45% coverage of the vesicle surface. It will be interesting to compare this estimate with EM analyses of the lattice of subunits formed in the assembled coat. But current EM images of the COPII coat show rather densely packed features[22] and lack the contrast of those produced in clathrin studies, which benefit from the open network of that lattice[2,22].

Domain organization of Sec23 and Sec24

The structural relatedness of the subunits was suggested in EM images showing the bow-tie complex as two heart-shaped halves that interact about a dyad[21]. The crystal structure reveals that the folds of the subunits are closely related, but that the sequences have diverged substantially so that the residual sequence identity is only ~14% according to a structure-based alignment (Fig. 2b), and alignments calculated using ClustalW[37] alone give somewhat spurious results. The alignment in Fig. 2b includes homologous regions of ~750 residues, and excludes the divergent 180 residues at the Sec24 N terminus, the initial 132 of which are disordered in the crystals, consistent with the low content of hydrophobic side chains and high content of glutamine and proline residues in this region.

The Sec23 and Sec24 polypeptides fold into five distinct domains (Fig. 2): a β-barrel, a zinc finger, an α/β vWA[38] or 'trunk' domain, an all-helical region and a carboxy-terminal domain that closely resembles a gelsolin module[39]. The β-barrel is formed from ~180 residues contributed by three polypeptide segments (see Fig. 2b). The strands of the barrel lie roughly parallel to the membrane such that one end of the barrel forms part of the concave inner surface of the coat and the other end part of the membrane-distal surface. The barrel is constructed from two apposed sheets: the six-stranded sheet β4-10-24-23-19-21 (numbered in the membrane distal-to-

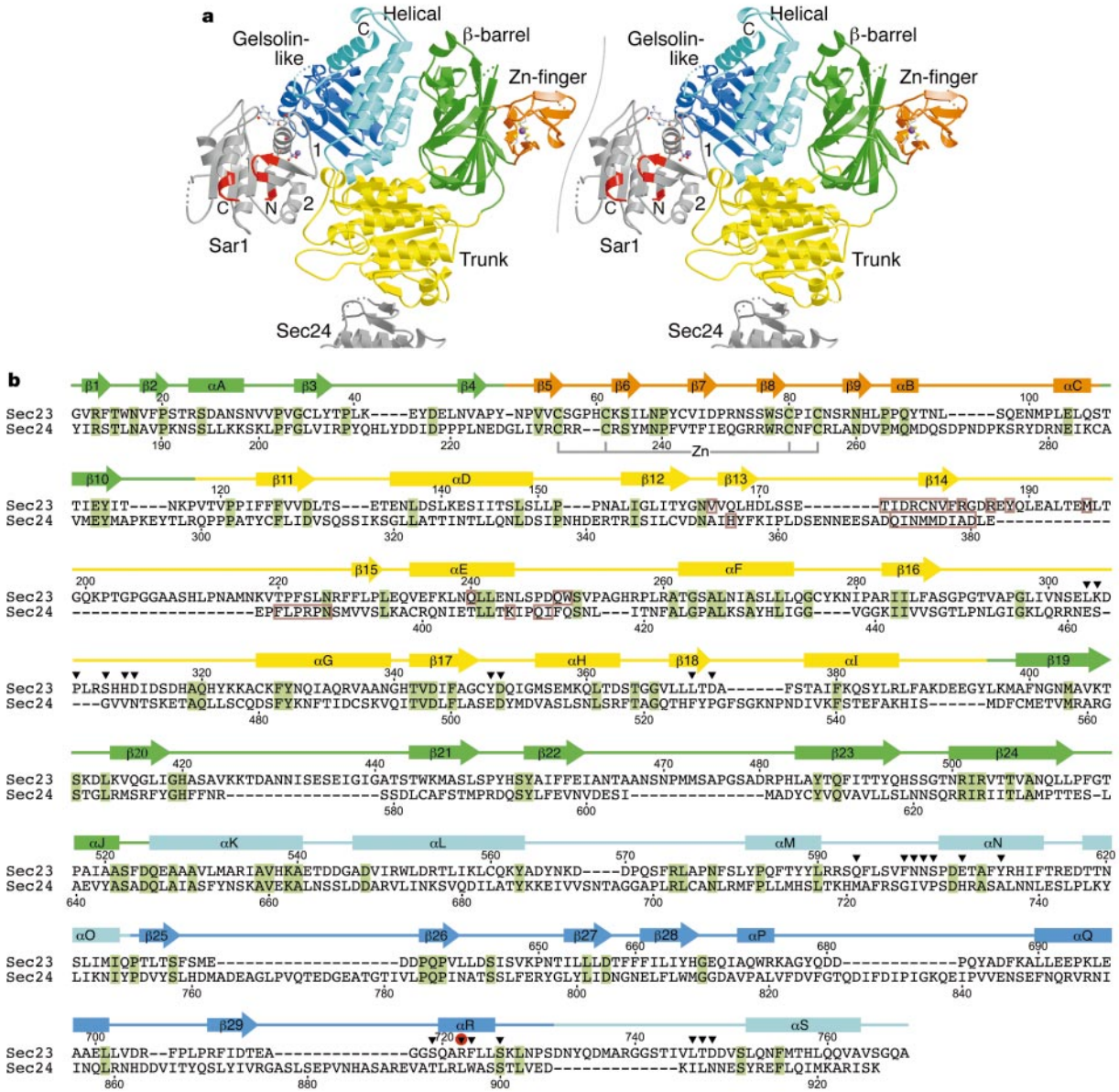

Figure 2 Domain structure and relatedness of Sec23 and Sec24. **a**, Stereo view, coloured according to the domain structure of Sec23, and oriented as in Fig. 1a. Sar1 and Sec24 are grey. The switch 1 and 2 regions of Sar1 are labelled 1 and 2, and the three membrane-proximal regions are red. **b**, Structure-based sequence alignment of *Saccharomyces cerevisiae* Sec23 and Sec24. The consensus secondary structure is shown above the sequences, coloured as in **a**, with α-helices depicted as cylinders and β-strands as arrows. Residues making intersubunit (between Sec23 and Sec24) interactions are outlined in brown, those residues of Sec23 contacting Sar1–GppNHp are denoted with black arrowheads and a red sphere for the catalytic arginine residue 722, and sequence identity is highlighted in green.

proximal direction) faces partly towards the zinc finger and partly towards solvent; sheet β1-3-22-20 faces the helical domain.

The ~55-residue zinc-finger domains lie against the β-barrels at the periphery of the complex (Figs 1a and 2a). The Zn-Cys_4 cluster is situated towards the membrane-proximal surface and several neighbouring arginine and lysine side chains on Sec23 and particularly Sec24 contribute to the basic inner surface (Fig. 3a). The trunk domain is formed from a single ~250-residue segment plugged into the barrel between strands β1 and β19. The trunk has an α/β fold with the vWA-type topology[38] and it forms the dimer interface, primarily involving strand β14 on Sec23 and Sec24 (Fig. 2); in addition, the trunk domain of Sec23 contacts Sar1. The α-helical domain forms from an ~105-residue segment with the C-terminal ~30 residues. Helices αK-L-M form a three-helix bundle with αL facing the concave inner surface. Three additional helices (N, O and S) complete the domain, with the αM–αN linker contacting Sar1. Finally, the gelsolin-like domain is formed by a contiguous stretch of 105 residues in Sec23, and a 150-residue region in Sec24 that is longer due to three insertions, all of which are disordered in the crystals. This is the only domain that does not contribute to the inner surface of the coat; instead it lies against the helical domain, forms critical catalytic interactions with Sar1 and the nucleotide, and extends radially away from the membrane, perhaps to form an interaction site for Sec13/31 (Figs 2a and 3a).

Heterodimer interface and Sec24 isoforms

In addition to the essential Sec23 and Sec24 proteins, yeast encode two (and mammals at least four) non-essential Sec24 homologues, Iss1 and Lst1, which are 55 and 23% identical to the Sec24 polypeptide; both form heterodimers with Sec23 (refs 36, 40). All three Sec24 isoforms interact with the ER-to-Golgi SNARE Sed5/syntaxin-5 (ref. 40), but only the Sec23/Lst1 pairing is proficient at packaging the plasma membrane ATPase Pma1 into vesicles[36]. Defining the interactions between these pre-budding complexes and sorting signals on SNARE and cargo molecules is a challenge for future studies, but at this stage an analysis of the conserved surface of Sec24 can suggest potential binding sites. In particular, the solvent-facing sheet of the β-barrel exposes a patch of highly conserved and hydrophobic side chains: Tyr 296, Val 557 and Leu 616 (Fig. 3b).

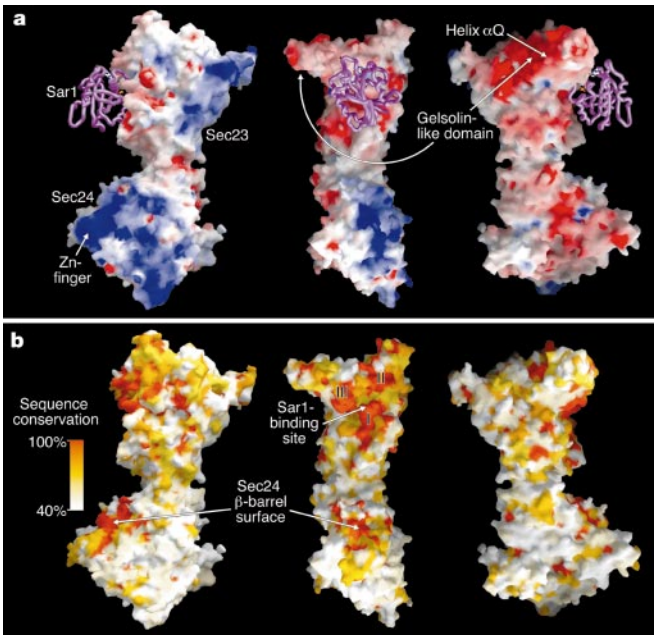

Figure 3 Surface features of the Sec23/24 complex **a**, Molecular surface coloured according to electrostatic potential[50]: negative potential is red and positive potential blue. Sar1 is drawn as a backbone tube in magenta. The three views are successive 90° rotations around a vertical axis. In the left image, the membrane-proximal surface faces forwards. **b**, Surface of Sec23/24 coloured according to sequence conservation of the underlying residues in an alignment of Sec23 sequences from six organisms (human Sec23A, *Drosophila melanogaster*, *Neurospora crassa*, *Schizosaccharomyces pombe*, *Arabidopsis thaliana* and *Saccharomyces cerevisiae*) and Sec24 sequences from four organisms (human Sec24A, *S. pombe*, *A. thaliana* and *S. cerevisiae*). Labels I, II and III indicate three highly conserved patches (see text for details). The orientations are the same as in **a**, with Sar1 omitted.

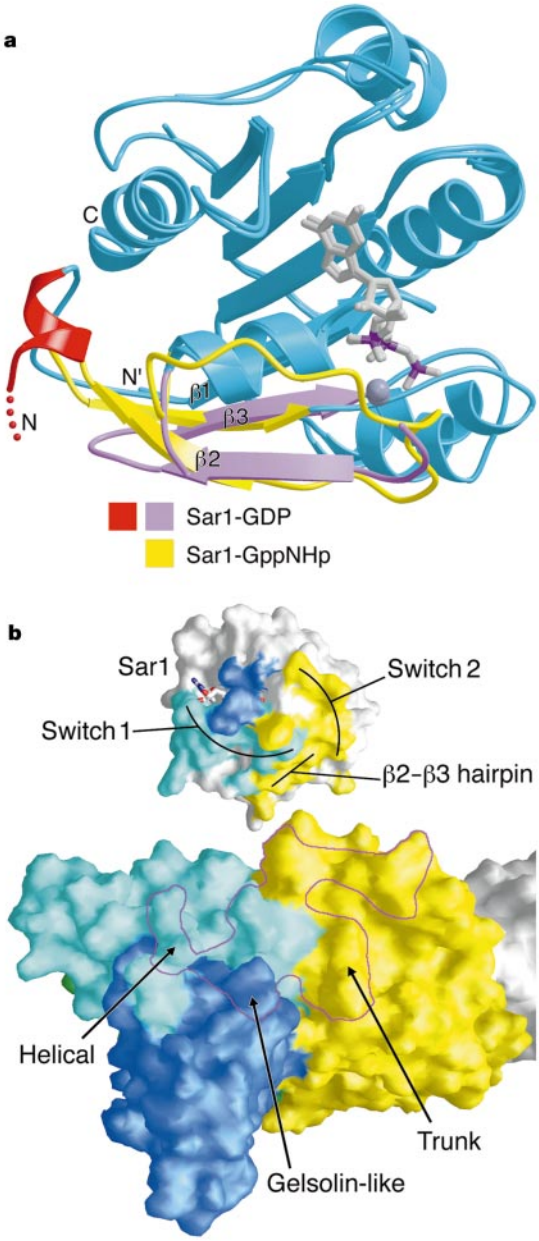

Figure 4 Conformational switching and protein–protein interactions. **a**, The nature of the Sar1 GDP–GTP conformational switch is suggested by superimposing mammalian Sar1–GDP[16] and yeast Sar1–GppNHp (from the Sec23–Sar1 complex). GTP-induced translation of β-strands 2 and 3 from the diphosphate (pink) to triphosphate (yellow) conformation eliminates the binding site for the N-terminal region (red) thought to interact with lipid membrane. Dark grey spheres are magnesium ions. N′ denotes the amino terminus (Gly 24) of the truncated Sar1 protein used in this study. **b**, 'Open book' view of the interfacial surfaces, with Sec23 and Sar1 rotated apart by 140°. The molecular surface of Sar1 is coloured to indicate the footprint of Sec23. The corresponding contact region on Sec23 is outlined in magenta. Colouring as in Fig. 2.

Sec23 recognizes Sec24 through an interface that buries ~900 Å^2 of Sec24 surface area. The contributing residues, chiefly from strand β14 and the β14–β15 loop, are not highly conserved among the yeast Sec24 isoforms, perhaps because strand β14 engages primarily in main-chain hydrogen bonds with the symmetry-related strand of Sec23, to create an inter-subunit β sheet joining the trunk domains (Figs 1 and 2). The exception is the β14–β15 loop region FLP (positions 385–387), which is conserved among the yeast Sec24 homologues and across species. Residues Phe 385 and Pro 387 from this region form van der Waals interactions with the universally conserved residues VFR (181–183) of Sec23 (Fig. 2b).

Sar1 conformational switch

Studies primarily on mammalian ARF1 have shown that ARF-family proteins possess a 'membrane anchor' that can be retracted in the GDP-bound state and exposed in the GTP state[3,18,20]. The anchor comprises an N-terminal amphipathic α-helix[17,34] that forms tight interactions with the membrane by inserting hydrophobic side chains into the bilayer[18]. In addition, the anchor on ARF proteins (although not on Sar1) is N-myristoylated, apparently not for the purpose of enhancing the membrane affinity of the GTP state but rather to increase the encounter frequency of ARF–GDP and membranes[41,42].

Biochemical knowledge of Sar1 is less advanced than that of ARF1, but the extent of sequence homology between these proteins (yeast Sar1 and human ARF1 are 33% identical) implies a common switching mechanism. And this is supported by the marked structural similarity of ARF1–GDP and Sar1–GDP, and of ARF1–GppNHp and Sar1–GppNHp (refs 16, 17, 19 and this study). Comparison of Sar1 in its GDP and GTP states provides a detailed view of the altered bonding around the γ phosphate that underlies the conformational switch, and this seems to be essentially the same as observed for ARF1, ARF6 and ARL2 (refs 19, 34, 35). For example, residue Thr 54 in the switch 1 region of Sar1–GppNHp (equivalent to Thr 35 in Ras) forms bonds to the γ phosphate and Mg^{2+}, and residue Gly 76 (Gly 60 in Ras) in switch 2 coordinates the γ phosphate.

Figure 4a shows the conformational change driven by these and associated interactions, and indicates how GTP binding may trigger the exposure of the Sar1 membrane anchor (red helix in Fig. 4a). In Sar1 proteins, an N-terminal extension of about ten residues replaces the myristoyl group and precedes the amphipathic α-helix. In Sar1–GDP, the helix is retracted into a surface pocket formed in part by the linker region of the β2–β3 hairpin[16]. The structure of Sar1–GppNHp shows that GTP binding causes an approximately 7 Å displacement of the hairpin—involving a change in register of the hydrogen bonding between strands β3 and β1—to eliminate the pocket for the anchor (Fig. 4a). This seems to explain how cytosol/membrane translocation of the protein can be controlled according to the type of nucleotide bound at its active site. And on the basis of these and analogous observations in the ARF1, ARF6 and ARL2 systems, it is probable that this mechanism is used for all membrane-dependent transactions in cells involving ARF-family proteins[19,34,35].

GTP-dependent coat recruitment

The arrangement of Sar1 in the pre-budding complex establishes how an activated ARF-family protein is oriented with respect to the membrane surface (Figs 1 and 2a). Figures 1b and 2a indicate that the β strands of Sar1 (β1-2-3 in particular) are approximately parallel to the membrane normal, such that three regions of Sar1 juxtapose the membrane: the anchor and additional N-terminal residues (in sum, residues 1–23, which are absent from the crystal structure), the linker region of the β2–β3 hairpin (residues 63–68), and C-terminal residues 187–190. Two basic residues, Lys 22 and Lys 68, which may contribute to membrane interaction, are located in these regions. More directly, *in vitro* studies on ARF1 have pinpointed interactions between acidic phospholipids and basic side chains in these regions, specifically the N-terminal residues Lys 15 and Lys 16, and C-terminal residues Arg 178 and Lys 181 (ref. 43; also note ARF1 residue Lys 59 in the β2–β3 linker).

The binding of GTP to Sar1 has two conformational consequences—the anchor is released for membrane interaction and the affinity for Sec23/24 is enhanced—the combined effect of which is membrane recruitment and formation of the pre-budding complex. Sec23 forms an extensive interface with Sar1–GppNHp (Fig. 4b) that encompasses ~1500 Å^2 of surface area, roughly 20% of the total Sar1 surface. Comparison with uncomplexed Sec23 shows that recruitment to Sar1 causes a subtle rotation of the barrel, zinc-finger, helical and gelsolin-like domains together as a rigid body with respect to the trunk domain, such that the curvature of Sec23/24 increases slightly, by about 1.5°. Sec23 recognizes the GTP state

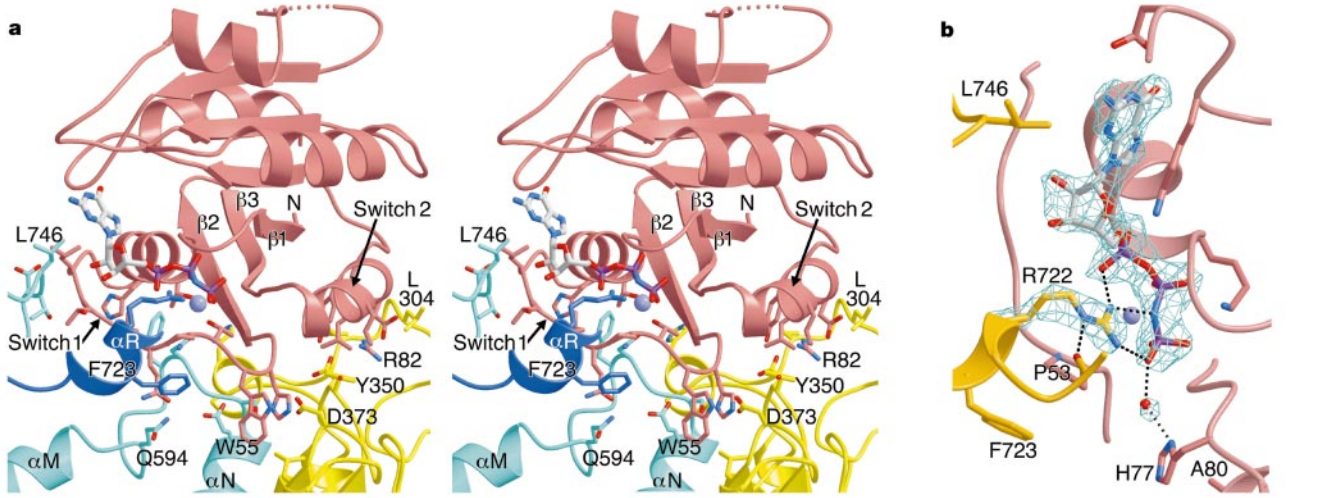

Figure 5 GTP-dependent coat recruitment and coat-dependent GTP hydrolysis **a**, Stereo view showing interactions between Sar1–GppNHp (pink) and Sec23 (coloured as in Fig. 2). Six amino-acid residues in Sec23 are labelled (triple-digit numbers), one from each discrete structural element that contributes to the interface. Sar1 residues W55 in switch 1 and R82 in switch 2 are labelled. Oxygen atoms are red, nitrogen atoms are blue, phosphorus atoms are magenta, and the magnesium ion is a blue-grey sphere. For clarity, portions of Sec23 have been removed from the foreground. **b**, Difference electron density (SA omit[49]) map at 2.5 Å resolution showing GppNHp, the catalytic residue Arg 722 of Sec23, and a water molecule (red sphere) positioned possibly for nucleophilic attack. Sar1 is pink and Sec23 is orange.

by contacting the switch 1, switch 2 and β2–β3 hairpin regions of Sar1 using residues from loops rather than secondary structure elements. Switch 1 is contacted primarily by the helical domain, and switch 2 and the hairpin by loops from the trunk domain that emanate from the carboxy ends of the β strands (Figs 2b, 4b and 5a). (Switch 1 is defined as Sar1 residues Lys 44 to Pro 57, and switch II as Asp 73 to Pro 89.)

The Sar1-binding site is the most conserved region of extended surface on Sec23/24 (Fig. 3b), and of the 24 residues contacting Sar1–GppNHp (calculated with a 3.8 Å cutoff), 12 are invariant among Sec23 sequences in the six organisms analysed (see legend to Fig. 3). These 12 cluster in three regions—labelled I, II and III in Fig 3b—of which patch I, formed by the αM–αN loop of the helical domain (Figs 2b and 5a), seems most critical for recognition of Sar1–GTP. The invariant residues FNNS (599–602) and Glu 605 of this loop contact switch 1 and the preceding α-helix (Fig. 5a). First, residue Phe 599 plugs into a hydrophobic pocket created by switch 1, and Asn 600 accepts a hydrogen bond from the side-chain OH of Thr 54, which also engages Mg^{2+} (but only in the GTP state). Second, Ser 602 and Glu 605 form interactions with switch 1 residues WHP (55–57), and these are reinforced by contacts involving trunk-domain residues from loops β17–αH and β18–αI (Figs 2b and 5a). The conservation of His 56 and Pro 57 in Sar1 proteins, and the systematic change from WHP (55–57) to IGF (49–51) in mammalian ARF1 and related isoforms suggests that this set of interactions may have a role in Sec23 discrimination by Sar1 and ARF proteins.

The interaction patch II (Fig. 3b) is formed by the αR–αS loop of the helical domain, which contacts residues RLA (47–49) of switch 1, and is notable owing to a van der Waals interaction between Leu 746 and the guanine base (Fig. 5). In general, the Sec23–Sar1 interface reveals a combination of hydrophobic and polar interactions, formed predominantly by side-chain–side-chain contacts, with few main-chain atoms involved. The final set of interactions, involving patch III, is formed between the gelsolin-like domain and the nucleotide-binding site of Sar1, and this may function not to recruit Sec23/24 in response to GTP binding but to control GTP hydrolysis and uncoating.

GTP hydrolysis and coat disassembly

Sar1–GTP, like ARF1–GTP, is catalytically inert. The problem of how GTP hydrolysis is triggered to cause COPII coat disassembly was resolved by the discovery that Sec23 is the Sar1 GAP[7], but this in turn raised the question of how the rate of COPII-mediated budding competes with the rate of COPII-mediated disassembly[8]. A 'two gear' mechanism has been proposed for the control of GTP hydrolysis, on the basis of kinetic studies showing that the lifetime of the Sec23/24–Sar1–GTP complex is relatively long (~30 s), and hydrolysis is accelerated about tenfold after Sec13/31 recruitment[8]. Hence, the COPII coat triggers its own disassembly, but the lifetime of GTP on Sar1 may be matched to the rate of budding to allow pre-budding complexes time to gather SNARE and cargo molecules, and completion of polymerization before disassembly[8].

Inspection of the Sec23–Sar1 interface reveals the mechanism of GTPase stimulation (Fig. 5). Sec23 inserts an arginine side chain, Arg 722, into the Sar1 active site to form bonds to the phosphates through its guanidinium group that will neutralize negative charges in the transition state. This type of mechanism, involving an arginine residue (arginine 'finger') supplied *in trans* to a G-protein active site, is a feature of many Ras-family proteins[44]. Evidently, isolated Sar1–GTP is inactive in part because it lacks the requisite catalytic machinery. Arg 722 and adjoining residues form helix αR on one face of the gelsolin-like domain (invariant patch III in Fig. 3b), and side-chain atoms of Ser 719, Arg 722, Phe 723 and Ser 726 directly contact Sar1. The side-chain OH of Ser 719 bonds to Asp 32 of the P-loop, but more importantly its C′O group forms a hydrogen bond to the catalytic Arg 722 to assist its positioning (Fig. 5b). Likewise, Phe 723 of Sec23 stabilizes the productive configuration at the active site through hydrophobic bonds to Pro 53 of switch 1, creating a sandwich of van der Waals contacts, the layers of which comprise Phe 723–Pro 53–Arg 722–P-loop (Fig. 5). This rigid arrangement may explain why Arg 722 forms a productive complex with the phosphates even in this ground-state structure, unlike in the Ras- and Rho-GAP systems where the arginine finger engages phosphates only in the transition state[44].

What can be inferred about the lack of GTPase activity of isolated Sar1–GTP? Studies on other GAP systems have shown that GAPs act in two ways to cause rate acceleration: by inserting an arginine residue, and by stabilizing switch 2 primarily to position a catalytic glutamine side chain (Gln 61 in Ras) and its associated nucleophilic water molecule[44]. In Sar1 proteins a histidine residue (His 77) replaces the glutamine, and in electron density maps His 77 bonds to an appropriately positioned water molecule (Fig. 5b). However, the position of switch 2 at the Sec23–Sar1 interface is unchanged from the position observed for the corresponding element in uncomplexed ARF1–GppNHp[19], suggesting that reconfiguration of switch 2 is not a principal feature of Sec23 action. Additional kinetic/mutagenetic studies are needed to assess what, if any, catalytic changes are caused by Sec23 in addition to Arg 722 insertion.

The lack of movement in switch 2 is at first sight surprising because the position it adopts is thought to contribute to ARF1 inactivity through misalignment of the catalytic glutamine, Gln 71 (ref. 19). Comparative analysis of the Sec23–Sar1, ARF1 and Ras–RasGAP active sites suggests that Sar1 proteins may avoid alterations to switch 2 by bonding the nucleophilic water in a new way. The uncommon substitution of the catalytic glutamine by histidine, and the change from QXXI (71–74) in ARF1 to HXXA (77–80) in Sar1 allows His 77 to lie against Ala 80 (Fig. 5b) and bond the nucleophilic water molecule in a manner not seen in other G-protein structures, through a configuration unfeasible with the bulkier isoleucine and longer glutamine side chains.

The mechanism by which Sec13/31 binding accelerates GTP hydrolysis tenfold[8] awaits further structural and biochemical analysis. A C-terminal region of the 140K Sec31 molecule, encompassing a proline-rich element, is thought to interact with both Sec23 and Sec24 (ref. 45). A potential binding site on Sec23 for Sec31 is the membrane-distal surface of the gelsolin-like domain (Fig. 3a), on the basis of two features: first, the radial extension of this domain away from the membrane gives it the appearance of a 'handle' for Sec31 binding; more directly, its close resemblance to a gelsolin module and the available gelsolin–actin complex structure[39] suggests an interaction site centred on helix αQ on the acidic outer surface (Fig. 3a). It will be interesting to discover the precise Sec31 binding site and whether rate acceleration involves direct or allosteric interaction with Sar1–GTP.

This analysis of the Sec23/24–Sar1 structure establishes the molecular basis for GTP-dependent recruitment and formation of a pre-budding complex, and for coat-dependent GTP hydrolysis that controls uncoating. And it provides the foundation for molecular-level studies of the mechanisms of SNARE and cargo selection by the COPII coat complex. □

Methods

Details of protein expression and purification are in Supplementary Information.

Crystallization

The Sec23/24 complex was crystallized at 22 °C by the hanging-drop method, by adding 1 μl of 35 mg ml^{-1} protein solution containing 15% (w/v) xylitol to 1 μl of 'mix' solution comprising 1.0 M ammonium formate, 18 mM dithiothreitol and 0.1 M HEPES buffer at pH 7.5, and placing this over well solution containing 1.9 M ammonium formate and 0.1 M HEPES at pH 7.5. The crystals are orthorhombic, space group $P2_12_12_1$ ($a = 90.3$ Å, $b = 126.4$ Å, $c = 180.2$ Å). Crystals were transferred to a cryoprotection solution of 6.0 M sodium formate, 10% xylitol and 0.05 M Hepes at pH 7.5, and flash-frozen in liquid propane. For crystallization of the Sec23–Sar1 complex, equimolar Sec23 and Sar1–GppNHp (prepared as described[19]) were mixed to give a 44 mg ml^{-1} protein solution. We

mixed 1 μl with 1 μl of well solution comprising 12% (w/v) PEG 1500, 10% (v/v) isopropanol, 0.2 M ammonium acetate and 50 mM MES buffer at pH 6.5. Orthorhombic ($P2_12_12_1$) crystals grow with two copies of the Sec23–Sar1 complex in the asymmetric unit (cell dimensions $a = 47.2$ Å, $b = 151.2$ Å, $c = 271.6$ Å). Before flash-freezing, crystals were cryoprotected by transfer to well solution containing additional 15% (v/v) ethylene glycol.

Structure determination

Diffraction data for Sec23/24 were measured at beamline X-25 of the National Synchrotron Light Source (NSLS). Data were processed with programs DENZO and SCALEPACK[46] (Supplementary Information Table 1). The structure of Sec23/24 was solved by molecular replacement with the program AMORE[48] using Sec23 as the search model. A translation function (8.0–4.0 Å resolution) gave a single unambiguous (20σ) peak. The model was improved using rigid-body and positional refinement, and a model for Sec24 was built in the resulting electron density map. Further refinement with CNS[49] yielded a final *R*-factor of 20.5% ($R_{free} = 25.3\%$) for data between 25.0 and 2.75 Å resolution. As is found in the Sec23/Sar1 structure, residues Val 30 and Phe 659 are the only Ramachandran plot outliers. The final model is complete with the exception of the disordered Sec23 residues 199–220, 466–480, 543–546 and 728–749, and Sec24 residues 1–132, 363–371, 463–466, 770–778, 823–847 and 876–887.

MAD data for the Sec23–Sar1 complex were collected at beamline F-2 of the Cornell High Energy Synchrotron Source (CHESS) (Supplementary Information Table 1). Data were processed as before, and MAD analysis was done with the program SOLVE[47] using data between 25.0 and 2.7 Å resolution. SOLVE found 23 of the 32 selenium sites plus two zinc sites, and yielded a mean figure of merit of 0.31 (0.17 for the highest-resolution bin). Density modification was carried out, including non-crystallographic symmetry averaging, with program DM (CCP4 suite[48]). This produced a high-quality electron density map, and a nearly complete model was built with the aid of the known structure of ARF1–GppNHp[19]. Refinement of atomic positions and temperature factors, together with a bulk solvent correction, with the program CNS[49] reduced the *R*-factor to a final value of 23.8% ($R_{free} = 29.5\%$) for data between 25.0 and 2.5 Å resolution. The model comprises 13,885 protein atoms, two molecules of Mg^{2+}–GppNHp, two zinc atoms and 171 water molecules. There are four Ramachandran plot outliers: residues Val 30 and Phe 659 in each copy of Sec23. Inspection of electron density maps confirms that these residues are modelled correctly, and adopt high-energy conformations owing to their location in tight turns. The following residues have not been modelled because of weak or no associated electron density: in each molecule of Sec23, residues 96–98, 204–218, 466–482, 732–742; and in each Sar1, residues 157–159.

Received 27 May; accepted 19 July 2002; doi:10.1038/nature01040.

1. Mellman, I. & Warren, G. The road taken: past and future foundations of membrane traffic. *Cell* **100,** 99–112 (2000).
2. Kirchhausen, T. Three ways to make a vesicle. *Nature Rev. Mol. Cell Biol.* **1,** 187–198 (2000).
3. Serafini, T. *et al.* ADP-ribosylation factor is a subunit of the coat of Golgi-derived COP-coated vesicles: a novel role for a GTP-binding protein. *Cell* **67,** 239–253 (1991).
4. Donaldson, J. G., Cassel, D., Kahn, R. A. & Klausner, R. D. ADP-ribosylation factor, a small GTP-binding protein, is required for binding of the coatomer protein β-COP to Golgi membranes. *Proc. Natl Acad. Sci. USA* **89,** 6408–6412 (1992).
5. Barlowe, C. *et al.* COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77,** 895–907 (1994).
6. Tanigawa, G. *et al.* Hydrolysis of bound GTP by ARF protein triggers uncoating of Golgi-derived COP-coated vesicles. *J. Cell Biol.* **123,** 1365–1371 (1993).
7. Yoshihisa, T., Barlowe, C. & Schekman, R. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic reticulum. *Science* **259,** 1466–1468 (1993).
8. Antonny, B., Madden, D., Hamamoto, S., Orci, L. & Schekman, R. Dynamics of the COPII coat with GTP and stable analogues. *Nature Cell Biol.* **3,** 531–537 (2001).
9. Nagahama, M. *et al.* A v-SNARE implicated in intra-Golgi transport. *J. Cell Biol.* **133,** 507–516 (1996).
10. Springer, S. & Schekman, R. Nucleation of COPII vesicular coat complex by endoplasmic reticulum to Golgi vesicle SNAREs. *Science* **281,** 698–700 (1998).
11. Grabowski, R. & Gallwitz, D. High-affinity binding of the yeast *cis*-Golgi t-SNARE, Sed5p, to wild-type and mutant Sly1p, a modulator of transport vesicle docking. *FEBS Lett.* **411,** 169–172 (1997).
12. Matsuoka, K. *et al.* COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell* **93,** 263–275 (1998).
13. Nakano, A., Brada, D. & Schekman, R. A membrane glycoprotein, Sec12p, required for protein transport from the endoplasmic reticulum to the Golgi apparatus in yeast. *J. Cell Biol.* **107,** 851–863 (1988).
14. Barlowe, C. & Schekman, R. SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. *Nature* **365,** 347–349 (1993).
15. Weissman, J. T., Plutner, H. & Balch, W. E. The mammalian guanine nucleotide exchange factor mSec12 is essential for activation of the Sar1 GTPase directing endoplasmic reticulum export. *Traffic* **2,** 465–475 (2001).
16. Huang, M. *et al.* Crystal structure of Sar1-GDP at 1.7 Å resolution and the role of the NH2 terminus in ER export. *J. Cell Biol.* **155,** 937–948 (2001).
17. Amor, J. C., Harrison, D. H., Kahn, R. A. & Ringe, D. Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature* **372,** 704–708 (1994).
18. Antonny, B., Beraud-Dufour, S., Chardin, P. & Chabre, M. N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange. *Biochemistry* **36,** 4675–4684 (1997).
19. Goldberg, J. Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell* **95,** 237–248 (1998).
20. Beraud-Dufour, S., Paris, S., Chabre, M. & Antonny, B. Dual interaction of ADP ribosylation factor 1 with Sec7 domain and with lipid membranes during catalysis of guanine nucleotide exchange. *J. Biol Chem.* **274,** 37629–37636 (1999).
21. Lederkremer, G. Z. *et al.* Structure of the Sec23p/24p and Sec13p/31p complexes of COPII. *Proc. Natl Acad. Sci. USA* **98,** 10704–10709 (2001).
22. Matsuoka, K., Schekman, R., Orci, L. & Heuser, J. E. Surface structure of the COPII-coated vesicle. *Proc. Natl Acad. Sci. USA* **98,** 13705–13709 (2001).
23. Kuehn, M. J., Herrmann, J. M. & Schekman, R. COPII-cargo interactions direct protein sorting into ER-derived transport vesicles. *Nature* **391,** 187–190 (1998).
24. Aridor, M., Weissman, J., Bannykh, S., Nuoffer, C. & Balch, W. E. Cargo selection by the COPII budding machinery during export from the ER. *J. Cell Biol.* **141,** 61–70 (1998).
25. Peng, R., Grabowski, R., De Antoni, A. & Gallwitz, D. Specific interaction of the yeast *cis*-Golgi syntaxin Sed5p and the coat protein complex II component Sec24p of endoplasmic reticulum-derived transport vesicles. *Proc. Natl Acad. Sci. USA* **96,** 3751–3756 (1999).
26. Votsmeier, C. & Gallwitz, D. An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. *EMBO J.* **20,** 6742–6750 (2001).
27. Martinez-Menarguez, J. A., Geuze, H. J., Slot, J. W. & Klumperman, J. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell* **98,** 81–90 (1999).
28. Balch, W. E., McCaffery, J. M., Plutner, H. & Farquhar, M. G. Vesicular stomatitis virus glycoprotein is sorted and concentrated during export from the endoplasmic reticulum. *Cell* **76,** 841–852 (1994).
29. ter Haar, E., Musacchio, A., Harrison, S. C. & Kirchhausen, T. Atomic structure of clathrin: a β propeller terminal domain joins an α zigzag linker. *Cell* **95,** 563–573 (1998).
30. Collins, B. M., McCoy, A. J., Kent, H. M., Evans, P. R. & Owen, D. J. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109,** 523–535 (2002).
31. Huang, M., Weissman, J. T., Wang, C., Balch, W. E. & Wilson, I. A. Protein engineering for crystallization of the GTPase Sar1 that regulates ER vesicle budding. *Acta Crystallogr. D* **58,** 700–703 (2002).
32. Paris, S. *et al.* Role of protein–phospholipid interactions in the activation of ARF1 by the guanine nucleotide exchange factor Arno. *J. Biol. Chem.* **272,** 22221–22226 (1997).
33. Goldberg, J. Structural and functional analysis of the ARF1–ARFGAP complex reveals a role for coatomer in GTP hydrolysis. *Cell* **96,** 893–902 (1999).
34. Pasqualato, S., Menetrey, J., Franco, M. & Cherfils, J. The structural GDP/GTP cycle of human Arf6. *EMBO Rep.* **2,** 234–238 (2001).
35. Hanzal-Bayer, M., Renault, L., Roversi, P., Wittinghofer, A. & Hillig, R. C. The complex of Arl2-GTP and PDEδ: from structure to function. *EMBO J.* **21,** 2095–2106 (2002).
36. Shimoni, Y. *et al.* Lst1p and Sec24p cooperate in sorting of the plasma membrane ATPase into COPII vesicles in *Saccharomyces cerevisiae*. *J. Cell Biol.* **151,** 973–984 (2000).
37. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680 (1994).
38. Ponting, C. P., Aravind, L., Schultz, J., Bork, P. & Koonin, E. V. Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. *J. Mol. Biol.* **289,** 729–745 (1999).
39. Robinson, R. C. *et al.* Domain movement in gelsolin: a calcium-activated switch. *Science* **286,** 1939–1942 (1999).
40. Peng, R., De Antoni, A. & Gallwitz, D. Evidence for overlapping and distinct functions in protein transport of coat protein Sec24p family members. *J. Biol. Chem.* **275,** 11521–11528 (2000).
41. Franco, M., Chardin, P., Chabre, M. & Paris, S. Myristoylation is not required for GTP-dependent binding of ADP-ribosylation factor ARF1 to phospholipids. *J. Biol. Chem.* **268,** 24531–24534 (1993).
42. Franco, M., Chardin, P., Chabre, M. & Paris, S. Myristoylation-facilitated binding of the G protein ARF1GDP to membrane phospholipids is required for its activation by a soluble nucleotide exchange factor. *J. Biol. Chem.* **271,** 1573–1578 (1996).
43. Randazzo, P. A. Functional interaction of ADP-ribosylation factor 1 with phosphatidylinositol 4,5-bisphosphate. *J. Biol. Chem.* **272,** 7688–7692 (1997).
44. Scheffzek, K., Ahmadian, M. R. & Wittinghofer, A. GTPase-activating proteins: helping hands to complement an active site. *Trends Biochem. Sci.* **23,** 257–262 (1998).
45. Shaywitz, D. A., Espenshade, P. J., Gimeno, R. E. & Kaiser, C. A. COPII subunit interactions in the assembly of the vesicle coat. *J. Biol. Chem.* **272,** 25413–25416 (1997).
46. Otwinoski, W. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
47. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
48. Collaborative Computational Project Number 4 The CCP4 suite: programs for X-ray crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
49. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
50. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).

Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com).

Acknowledgements

We thank E. Mossessova for the Sar1 expression construct, J. Walker for assistance and advice on data collection and processing, C. Heaton for use of synchrotron facilities at CHESS, and M. Becker at NSLS. This work was supported by grants from the National Institutes of Health, the Howard Hughes Medial Institute, and the Pew Scholars Program in the Biomedical Sciences.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.G. (e-mail: jonathan@ximpact4.ski.mskcc.org). Coordinates have been deposited with the Protein Data Bank under accession codes 1M2O and 1M2V.