# TANGO1/cTAGE5 receptor as a polyvalent template for assembly of large COPII coats

Wenfu Ma[a,b] and Jonathan Goldberg[a,b,1]

[a]Howard Hughes Medical Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065; and [b]Structural Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065

Edited by David J. Owen, University of Cambridge, Cambridge, United Kingdom, and accepted by Editorial Board Member Pietro De Camilli July 9, 2016 (received for review April 12, 2016)

**The supramolecular cargo procollagen is loaded into coat protein complex II (COPII)-coated carriers at endoplasmic reticulum (ER) exit sites by the receptor molecule TANGO1/cTAGE5. Electron microscopy studies have identified a tubular carrier of suitable dimensions that is molded by a distinctive helical array of the COPII inner coat protein Sec23/24•Sar1; the helical arrangement is absent from canonical COPII-coated small vesicles. In this study, we combined X-ray crystallographic and biochemical analysis to characterize the association of TANGO1/cTAGE5 with COPII proteins. The affinity for Sec23 is concentrated in the proline-rich domains (PRDs) of TANGO1 and cTAGE5, but Sec23 recognizes merely a PPP motif. The PRDs contain repeated PPP motifs separated by proline-rich linkers, so a single TANGO1/cTAGE5 receptor can bind multiple copies of coat protein in a close-packed array. We propose that TANGO1/cTAGE5 promotes the accretion of inner coat proteins to the helical lattice. Furthermore, we show that PPP motifs in the outer coat protein Sec31 also bind to Sec23, suggesting that stepwise COPII coat assembly will ultimately displace TANGO1/cTAGE5 and compartmentalize its operation to the base of the growing COPII tubule.**

vesicle traffic | coat protein | procollagen

The coat protein complex II (COPII)-coated vesicles transport secretory and plasma membrane proteins from the endoplasmic reticulum (ER). COPII vesicle budding involves a stepwise assembly reaction: Membrane-bound Sar1-GTP recruits Sec23/24 to form the inner coat complex that, in turn, recruits the Sec13/31 outer coat protein (1). Sec13/31 self-assembles into a polyhedron, and in the process, it sculpts the ER membrane into a bud, producing a COPII-coated vesicle with a diameter of ~60 nm (2–4).

Small cargo molecules are captured during the budding reaction through mechanisms that are well established (5). However, the formation of canonical 60-nm COPII vesicles cannot explain the packaging of large cargos, such as the procollagen fibril with a length of 300–400 nm. Procollagen follows the conventional route of secretion taken by other extracellular proteins, and its ER export evidently depends on COPII carriers because mutations in genes encoding COPII proteins result in collagen secretion blockade, impaired extracellular matrix deposition, and abnormal craniofacial development (6–8). A key discovery was the identification of the receptor TANGO1 and its coreceptor cTAGE5 as ER-localized machinery for loading procollagen into COPII carriers (9, 10). TANGO1 has a luminal SH3 domain shown to interact with collagen VII and a cytosolic domain that interacts with Sec23/24. Both TANGO1 and cTAGE5 are required for collagen export from the ER (9, 10). The TANGO1 knockout mouse has generalized defects in extracellular matrix formation, owing to a block in ER export of multiple collagen types (11).

Recent research has implicated post-ER membranes in TANGO1-mediated carrier formation (12, 13); however, the mechanism of action of TANGO1/cTAGE5 in procollagen export remains unclear. The possibility that it acts as a receptor linking luminal procollagen cargo to cytosolic COPII is challenged by the observation that TANGO1 does not depart in the COPII carrier (9). Alternatively, TANGO1/cTAGE5 might constitute machinery that switches COPII coat assembly from small spherical vesicles to large carriers, but such a mechanism has not been explored to date. Important in this respect is the observation of tubular COPII carriers with dimensions commensurate with procollagen cargo; such tubules form, along with small vesicles, in budding reactions in vivo and in vitro (14–16). COPII-coated tubules are observed as straight-sided tubes (17), or with a beads-on-a-string appearance in which regular constrictions may arise from aborted attempts at membrane fission (14, 16). Whereas vesicles contain a Sec23/24 inner coat and a self-assembled Sec13/31 cage with isometric (cubic) symmetry, tubules are coated with a distinctive, close-packed Sec23/24 helical lattice and a rhomboidal Sec13/31 cage (17). The realization that COPII coat proteins can form two distinct and mutually incompatible lattices suggests the hypothesis that COPII carrier formation involves two self-assembly reactions that compete to determine the shape of the COPII carrier—either the isometric small vesicle or the helical tubule—with the outcome of budding influenced by regulatory proteins that promote one or other lattice.

In this study, we have characterized the interaction of TANGO1/cTAGE5 with COPII proteins. We report that TANGO1/cTAGE5 binds to the inner coat protein Sec23 via a simple PPP motif. The flexible cytosolic regions of TANGO1 and cTAGE5 molecules contain repeated PPP motifs separated by proline-rich linkers, so a single TANGO1/cTAGE5 receptor can bind multiple copies of coat protein in a close-packed array. From these results, we propose a mechanism for TANGO1/cTAGE5-assisted assembly of large COPII carriers.

## Significance

**Proteins destined for secretion from cells enter the secretory pathway in coat protein complex II (COPII)-coated vesicles budding from the endoplasmic reticulum (ER). Most cargo proteins are small and exit the ER in 60-nm vesicles. However, some secretory cargos are too large to enter such carriers; in particular, the procollagen precursor of the extracellular matrix exits the ER as a 300- to 400-nm fibril. Recent research suggests that procollagen may be packaged into large COPII-coated tubules, guided by the receptor molecule TANGO1/cTAGE5. We show that each TANGO1/cTAGE5 receptor protein has a multiplicity of binding sites to recruit and concentrate COPII proteins. We propose the model that TANGO1/cTAGE5 instructs the COPII coat to form large tubular carriers.**

Author contributions: J.G. designed research; W.M. performed research; W.M. and J.G. analyzed data; and J.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.J.O. is a Guest Editor invited by the Editorial Board.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5KYN, 5KYX, 5KYU, 5KYW, and 5KYY).

[1]To whom correspondence should be addressed. Email: goldberj@mskcc.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1605916113/-/DCSupplemental.

## Results

**Interaction of TANGO1 and cTAGE5 with COPII Proteins.** Interactions between TANGO1/cTAGE5 and the COPII inner coat protein Sec23/24 were identified by yeast two-hybrid analysis in two published studies; the homologous proline-rich domains (PRDs) of TANGO1 and cTAGE5 were both shown to interact with human Sec23a (9, 10). The PRD of TANGO1 is also required for ER exit site localization, presumably by virtue of its interaction with Sec23/24 (9). The schematic in Fig. 1*A* indicates the position of the PRD at the C terminus of the TANGO1 and cTAGE5 polypeptides.

To map the Sec23/24 binding site to a circumscribed region of TANGO1, we prepared a series of polypeptides (Fig. 1*A*, *Lower*) and tested these for binding to purified Sec23 and Sec24 proteins (*Materials and Methods*). TANGO1 polypeptides, expressed as C-terminal fusions to maltose binding protein (MBP), were immobilized on amylose resin beads and probed for binding to human Sec23a. Sec23a bound to a construct encompassing the entire PRD plus coiled coil 2 of TANGO1, residues 1444–1907 (Fig. 1*B*, lane 3). No appreciable binding was observed between Sec23a and MBP alone (lane 2). The affinity for Sec23 was contained in the C-terminal half of the TANGO1 PRD (lane 5), and no significant binding was observed to the N-terminal half (lane 4). We repeated this analysis with cTAGE5 (Fig. 1*C*), and found that cTAGE5 residues 651–804 bound to Sec23a (Fig. 1*C*, lane 4), as did shorter cTAGE5 fragments (lanes 5 and 6). These data are consistent with two-hybrid interactions reported (9, 10). In addition to Sec23a, the paralog Sec23b bound to TANGO1 and cTAGE5 (Fig. 1*D*). Finally, we tested for interactions with the inner coat protein Sec24c; neither TANGO1 nor cTAGE5 PRD bound to Sec24c, because binding was at background levels (Fig. 1*E*). This finding contrasts with experiments that reported a two-hybrid interaction with Sec24c, albeit significantly weaker than Sec23a interactions with TANGO1 and cTAGE5 (10). Taken together, we suggest that TANGO1/cTAGE5 interacts specifically with the Sec23 component of the COPII inner coat complex.

Next, we tested the series of short TANGO1 polypeptides (Fig. 1*A*) to pinpoint the interaction with Sec23a and to define a complex for crystallographic analysis. The tight-binding TANGO1 construct 1751–1907 was used as a starting point (Fig. 1*B*, lane 5). N-terminal truncation of this construct reduced binding to Sec23a (Fig. 1*B*, compare, for example, lanes 7 and 8), in particular beyond residue 1780 (compare lanes 10 and 12). Likewise, C-terminal truncations also reduced binding (Fig. 1*B*, compare, for example, lanes 6 and 12). Because the TANGO1 PRD is an intrinsically disordered polypeptide, a clear cutoff effect might not be expected from these experiments; nevertheless, it was puzzling to find that affinity for Sec23a depends on TANGO1 residues that are dispersed along 157 residues of polypeptide.

The biochemical analysis indicated that TANGO1 residues 1780–1840 retained appreciable binding to Sec23a (Fig. 1*B*, lane 11), albeit weaker than longer forms, so we chose this polypeptide for a crystallographic analysis of the complex with Sec23.

**Molecular Recognition of TANGO1 by Sec23.** The complex comprising full-length human Sec23a and TANGO1 residues 1780–1840 was cocrystallized, and the structure was determined and refined to 2.6 Å resolution (Table S1). Despite the fact that 61 residues of TANGO1 are present in the crystals, difference electron density reveals just a three-residue peptide element bound to a site on the gelsolin-like domain of Sec23a (Fig. 2 *A* and *B*) [Sec23 domain nomenclature was defined previously (18)]. The tripeptide can be best modeled as three proline residues in the polyproline type II (PPII) helical conformation. The central proline in particular is positioned between Sec23a residues Trp667 and Tyr678 (Fig. S1, *Left*). Indeed, the entire binding site is formed from aromatic residues, and a striking similarity is observed with the polyproline-binding sites of evolutionarily unrelated domains. Fig. S1 compares

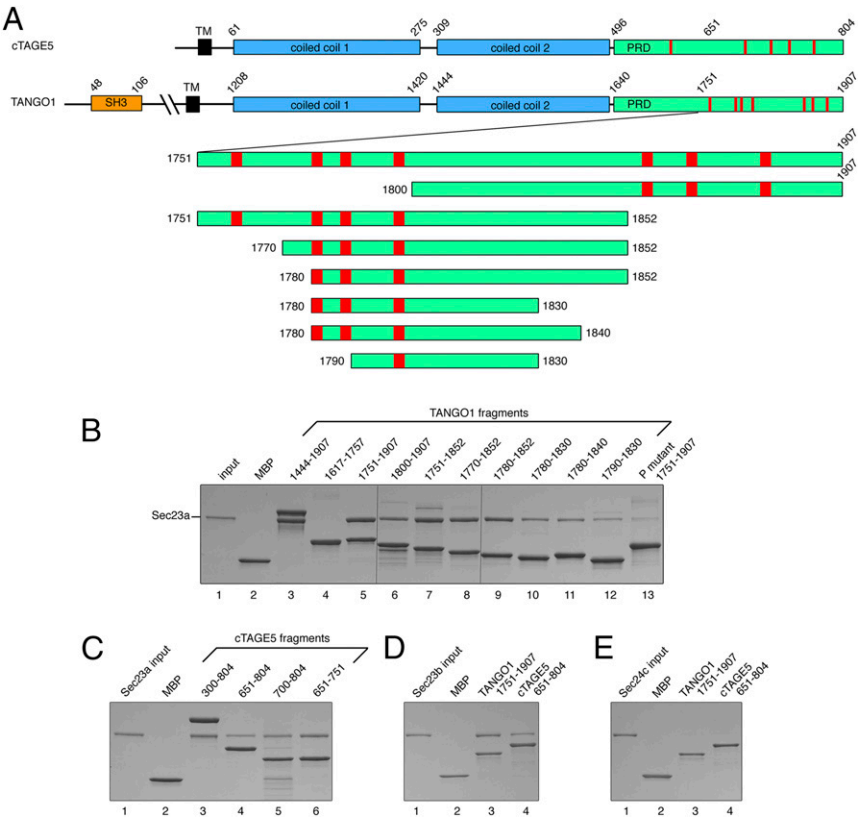


**Fig. 1.** Dissection of TANGO1 and identification of Sec23-binding regions. (*A*) *Upper* diagram shows the domain organization of the homologous human cTAGE5 and TANGO1 proteins. Cytosolic regions are drawn to the right of the transmembrane domain. The PRDs of TANGO1 and cTAGE5 are colored green, and the repeated PPP sequences that our experiments identify as the Sec23-binding motif are indicated in red. TANGO1 and cTAGE5 interact via their second (C-terminal) coiled coil domains (10). SH3 denotes the luminal Src homology 3 domain of TANGO1. In addition to the single transmembrane domain, TANGO1 has an adjacent membrane-binding sequence that penetrates but does not cross the bilayer (9). The lower diagram shows the series of constructs of the TANGO1 PRD that were prepared to explore the interaction with Sec23/24. (*B*) TANGO1 polypeptides fused to MBP were immobilized on amylose resin beads, which were then incubated with a mixture containing 0.8 mg/mL human Sec23a. Proteins were analyzed by 4–20% (wt/vol) SDS/PAGE and Coomassie blue staining. (*C*) Sec23a also binds to the PRD of cTAGE5. (*D*) TANGO1 and cTAGE5 both bind to the alternative isoform Sec23b. (*E*) No binding is detected between human Sec24c and the PRDs of TANGO1 or cTAGE5.

Sec23 to the SH3 and GYF domains (WW and EVH1 domains offer additional examples). The domain topologies are unrelated, but all of the polyproline-binding sites are formed from aromatic residues positioned to form a series of ridges and grooves that accommodate the multiple proline side chains (19, 20).

The unexpected crystallographic result of a minimal PPP sequence of TANGO1 bound to Sec23a suggests a model of TANGO1 binding to Sec23 via a 3-aa motif that repeats through the PRD, thus providing a plausible explanation for our inability to pinpoint the binding sequence of TANGO1 in biochemical experiments: Seven copies of the PPP motif are present in the human TANGO1 PRD (cTAGE5 PRD contains five PPP motifs), and the crystallized fragment of TANGO1 (1780–1840) has three PPP motifs (Fig. 1*A*). However, the crystallographic result does not definitively assign the binding motif as being uniquely PPP, because the electron density is almost certainly a composite of multiple proline-rich tripeptide sequences from TANGO1 residues 1780–1840 and may include XPP or PPX sequences. Thus, we proceeded to test the preference of Sec23 for PPP and related motifs, and also to verify that the gelsolin-like domain is the binding site for TANGO1.

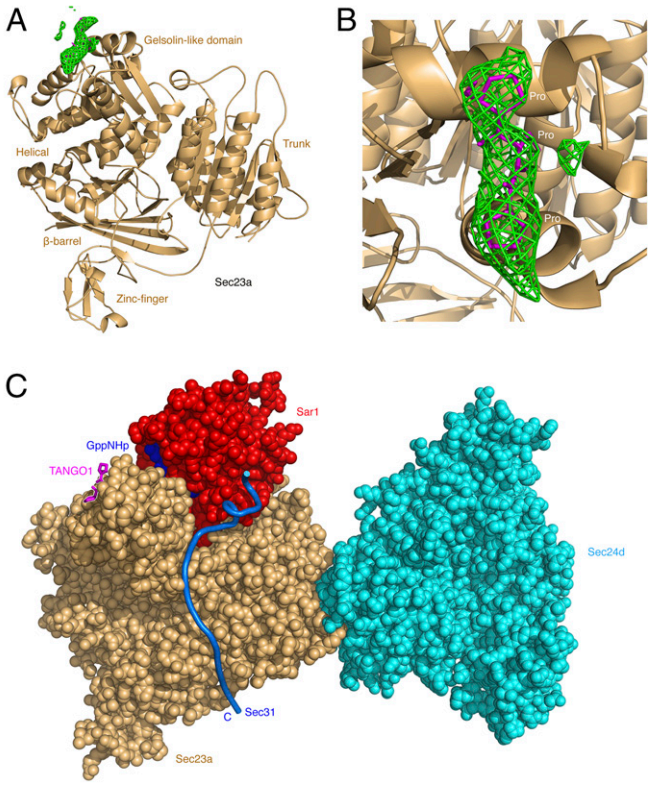


**Fig. 2.** TANGO1 binding site located by X-ray crystallography. (*A*) Crystal structure of Sec23a bound to TANGO1 (residues 1780–1840) is drawn in ribbon representation. Green contour lines show residual electron density (Fo-Fc map; 2.6 Å resolution, 2.3 σ contour level) following refinement of Sec23a coordinates, in the vicinity of the gelsolin-like domain. (*B*) Close-up view of the difference electron density for the bound TANGO1 peptide, with the density calculated as in *A*. A tripeptide PPP sequence is modeled in a PPII helical conformation. (*C*) Composite model to indicate the location of the TANGO1-binding site in relation to the complete COPII inner coat complex. This is a view toward the membrane-distal surface of the complex. Sec23a (colored gold) and TANGO1 PPP motif (magenta) are drawn in the same orientation as *A*. In addition, we have included the structure of Sec24d, colored cyan (data from ref. 21; PDB ID code 3EG9), as well as Sar1 (colored red with bound GppNHp nucleotide in blue) and the 50-residue binding element of Sec31, drawn as a dark blue worm (from the yeast crystal structure of Sec23•Sar1 complexed with Sec31, PDB ID code 2QTV) (22).

**Individual PPP Motifs Bind to Sec23.** To confirm that TANGO1 binds to the gelsolin-like domain of Sec23, we generated two Sec23a mutants (a double mutant Y672K/Y678A and a double mutant F628A/F681A), thus altering key aromatic residues (Fig. S1, *Left*). Consistent with the crystallographic observation, both mutants showed impaired ability to bind TANGO1 (Fig. 3 *A* and *B*). The tyrosine double-mutant Sec23a lost rather more affinity, possibly because of the critical role of residue Tyr678 in sandwiching the central proline residue of PPP (Fig. S1).

Next, we tested the binding of individual PPP motifs to Sec23a. A single PPP motif binds to Sec23 too weakly to measure by using conventional methods (e.g., isothermal calorimetry), because the binding signal cannot be distinguished from background with confidence. Our solution to this problem was to use crystallography to provide spatial information that confirms specific binding. We grew a crystal form of Sec23a/24d (described in ref. 21) in which the gelsolin-binding domain is unobstructed by crystal contacts, so that short proline-rich peptides could be soaked into, and equilibrated with, the binding site. Residual electron density at the gelsolin-binding domain was then assessed for evidence of specific binding. Four TANGO1 peptides, 8–10 residues in length, were synthesized. The sequences are partially overlapping (encompassing TANGO1 residues 1790–1816) and were chosen such that two peptides contain a PPP motif, one peptide has a PP motif, and the final peptide contains only three isolated proline residues (Fig. 3 *D*–*G*). Peptides were soaked at high concentrations (6 mM) into crystals of Sec23a/24d. The crystals diffract to only medium resolution (3.2–3.5 Å) (Table S1); nevertheless, specific binding could be established straightforwardly by the appearance of residual electron density. TANGO1 peptides GPRPLPPP and LPPPFGPGM (Fig. 3 *E* and *F*) show clear binding to Sec23a (compare the electron density with the location of TANGO1 PPP in Fig. 3*C*, which is oriented similarly). By contrast, TANGO1 peptides lacking the PPP motif—QLCGPFGPRP and GMRPPLGLRE in Fig. 3 *D* and *G*—do not bind to Sec23a.

These data indicate that TANGO1 binds to the gelsolin-like domain of Sec23a via a PPP motif that repeats through the TANGO1 and cTAGE5 cytosolic regions. In Fig. 2*C*, we summarize these findings in a composite atomic model to indicate the location of the bound PPP motif on the COPII inner coat complex, near to Sar1-GTP and in proximity to the "active fragment" (the Sec23•Sar1-binding element) of Sec31.

**Repeated PPP Motifs in TANGO1 and cTAGE5.** A region of the TANGO1 PRD containing seven PPP motifs, residues 1751–1907, seems to be responsible for all of the Sec23a-binding affinity, and forms of this polypeptide lacking the full complement of motifs bind less Sec23a (Fig. 1*B*). These results conform to our model of TANGO1 binding to Sec23 through an array of distributed PPP motifs; however, the experiment involves TANGO1 constructs of diverse sizes and compositions, so binding cannot easily be compared quantitatively. We measured, on a more systematic basis, how changes in the number of PPP repeats affect binding of TANGO1, using a series of TANGO1 polypeptides (residues 1751–1907) in which sets of proline residues were mutated sequentially to glycine or serine (Fig. S2; see Fig. S3 for sequences of the mutants, named M1 to M6).

A form of TANGO1 1751–1907 with all 46 of its proline residues mutated retains only residual affinity for Sec23a (mutant M6 in Fig. S2). Weak binding is observed between Sec23a and TANGO1 mutant M5, which has just two PPP motifs. Inclusion of an additional repeat (mutant M4 with three PPP motifs) increases binding further; each addition of PPP motifs increments the binding to Sec23a. These results show a clear correlation between the number of PPP repeats and the amount of Sec23a bound.

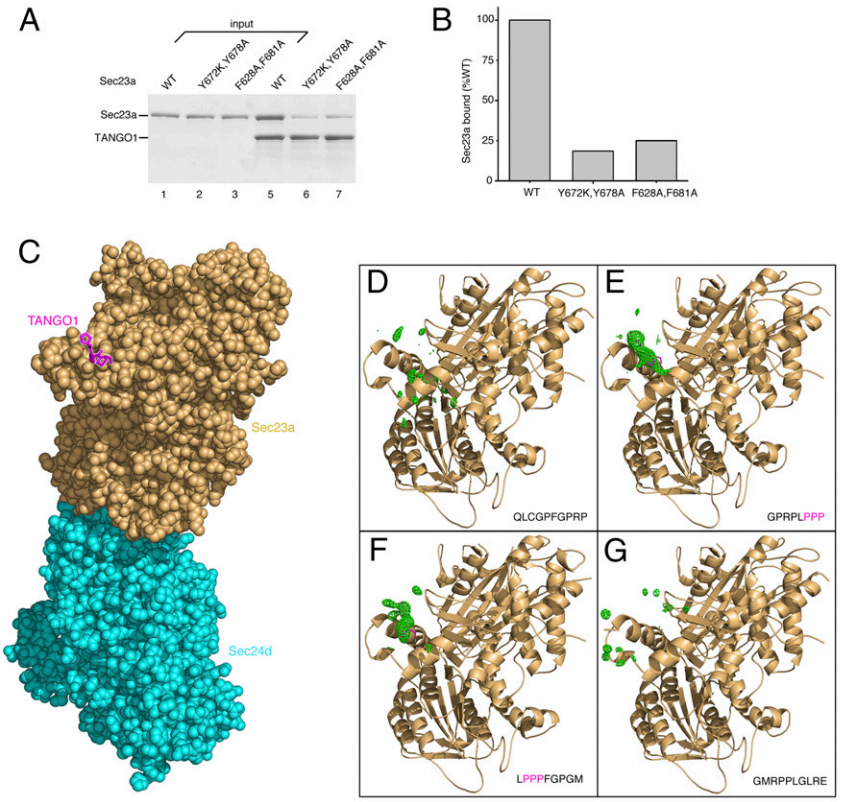


**Fig. 3.** Examination of the TANGO1-binding site on Sec23a, and of the Sec23-binding motif on TANGO1. (*A*) Mutations to Sec23a were introduced into the binding site defined crystallographically, and these are found to substantially reduce binding to TANGO1 PRD residues 1751–1907. Two mutants of Sec23a were tested, both double mutants (Y672K/Y678A and F628A/F681A); the location of these residues is shown in Fig. S1. (*B*) The reduction in binding to TANGO1 caused by mutation of Sec23a was quantified by densitometry of bands in *A*. (*C*) All-atom representation of the Sec23a/24d crystal structure with bound TANGO PPP peptide (peptide sequence GPRPLPPP) (Table S1). (*D*) Residual electron density in the vicinity of the gelsolin-like domain, following refinement of Sec23a/Sec24d crystals soaked with the TANGO1 peptide $^{1790}$QLCGPFGPRP$^{1799}$ (green contour lines show $F_o$–$F_c$ difference electron density at the 2.3 σ contour level, calculated at 3.4 Å resolution). Note that the picture is oriented as in *C*. The crystallographic result indicates no binding for this peptide. (*E*) Residual electron density in the vicinity of the gelsolin-like domain indicates bound TANGO1 peptide $^{1796}$GPRPLPPP$^{1803}$ ($F_o$–$F_c$ map; 3.5 Å resolution, 2.3 σ contour level). (*F*) Electron density shows bound TANGO1 peptide $^{1800}$LPPPFGPGM$^{1808}$ ($F_o$–$F_c$ map; 3.3 Å resolution, 2.3 σ contour level). (*G*) Absence of bound peptide at the gelsolin-like domain in crystals soaked with TANGO1 peptide $^{1807}$GMRPPLGLRE$^{1816}$ ($F_o$–$F_c$ map; 3.2 Å resolution, 2.3 σ contour level).

**PPP Motifs in the Sec31 Outer Coat Protein Bind to Sec23.** The COPII outer coat protein Sec31 contains an intrinsically disordered PRD that includes the active fragment for binding Sec23•Sar1 (22). Besides the repeated PPP motifs in TANGO1 and cTAGE5, we also noted the presence of conserved PPP motifs in this strategically important region of the Sec31 polypeptide. Specifically, in human Sec31a, PPP motifs are located at residues $^{838}$PPPP$^{841}$ and $^{944}$PPPP$^{947}$; the active fragment is located at residues 980–1015.

To test whether PPP motifs on Sec31 can bind to the gelsolin-like domain of Sec23, we generated a form of human Sec31a containing both PPP motifs plus the active fragment (in total, residues 835–1026), and expressed this as a C-terminal fusion to maltose-binding protein. Sec31a bound to Sec23a (Fig. 4*A*, lane 7); binding is weak because Sec31a has only two PPP motifs (compare, for example, with mutant M5 in Fig. S2). Importantly, Sec31a did not bind the Sec23a mutants Y672K/Y678A or F628A/F681A, establishing that the interaction was via the gelsolin-like domain (Fig. 4 *A* and *B*). The background binding to the tyrosine double mutant Sec23a was low, in concordance with the binding behavior of the two mutants toward TANGO1 (compare Fig. 4*B* with Fig. 3*B*).

It is important to note that, in this experiment, there is no evidence for Sec31a binding to Sec23a via the active fragment; this interaction depends on the presence of GTP-bound Sar1 (22). We verified this point by testing the binding of Sec31a (residues 835–1026) to Sec23a in the presence of Sar1 complexed with the nonhydrolyzable GTP analog GppNHp (we used soluble human Sar1a, residues 25–198, lacking the N-terminal membrane anchor). Indeed, the interaction highly depended on Sar1-GppNHp, which increased ∼fivefold the quantity of Sec23a bound to Sec31a (Fig. 4*C*, lanes 5 and 6). By contrast, binding of Sec23a to TANGO1 is independent of Sar1-GppNHp (lanes 7 and 8). Note that Sec23a binding to TANGO1 is appreciable in this experiment because we used construct 1751–1907 containing seven PPP motifs, whereas the Sec31 polypeptide has just two PPP motifs. If we adopt the data in Fig. S2 as a metric, it is clear that the interaction of Sec31a with Sec23a•Sar1 is significantly tighter than the binding of a single PPP motif to Sec23a.

Finally, we tested whether Sec31 can compete with a TANGO1 polypeptide for binding to Sec23 in vitro (Fig. S4). A TANGO1 polypeptide containing three PPP motifs (residues 1800–1907, fused to MBP) was immobilized on amylose resin beads and probed for binding to Sec23a in the presence (Fig. S4, lane 9) and absence (lane 8) of Sec31a (residues 835–1026) (*SI Materials and Methods*). A twofold molar excess of Sec31a relative to Sec23a is sufficient to markedly reduce the binding of Sec23a to TANGO1. Note the reduction in binding of both the Sar1-GppNHp and Sec23a proteins. The data support our model of a competition between the TANGO1 and Sec31 polypeptides for binding to Sec23.

These results, regarding Sec31 interaction with Sec23•Sar1, are summarized in the composite molecular model in Fig. 4*D*.

## Discussion

In this study, we have characterized the association between COPII coat proteins and the TANGO1/cTAGE5 receptor for procollagen packaging. We report that an unexpectedly short PPP motif in the TANGO1 polypeptide binds to Sec23a, and that TANGO1 and cTAGE5 contain repeated PPP motifs distributed across their flexible cytosolic PRD regions. This unusual arrangement may enable TANGO1/cTAGE5 to engage multiple copies of Sec23/24•Sar1 at ER exit sites, leading to the speculative model that the role of the receptor is to propagate the COPII inner coat lattice of a growing tubular carrier.

**Architecture of the COPII Inner Coat Complex.** The picture in Fig. 2*C* is a view toward the membrane-distal surface of the inner coat complex, to which the PPP motif and the Sec31 active fragment bind. The gelsolin-like domain of Sec23 extends radially away from the membrane, such that the PPP motif will reside ∼45 Å from the

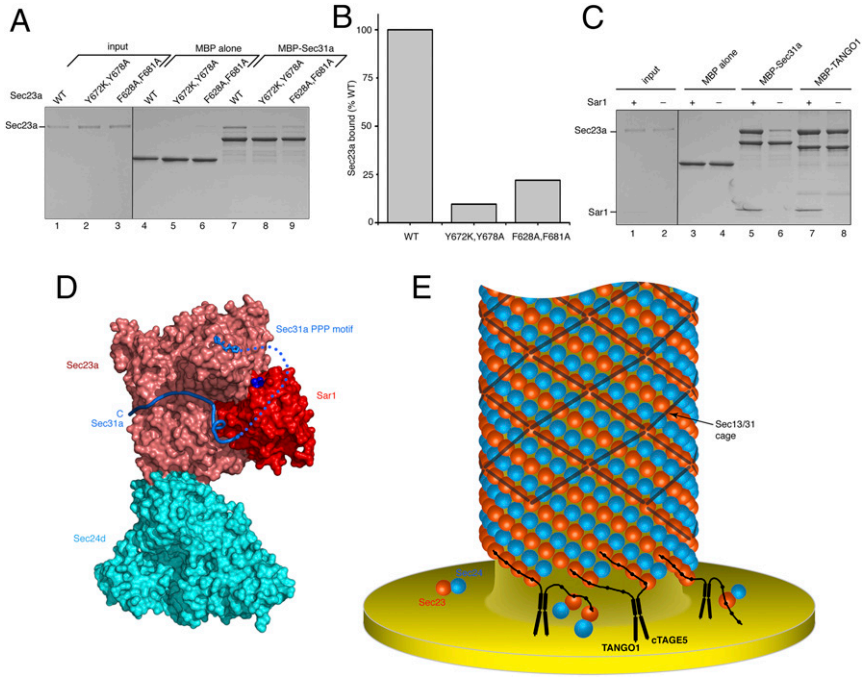


**Fig. 4.** Binding sites for Sec31 and TANGO1 on Sec23a•Sar1. (*A*) The polypeptide region of human Sec31 encompassing two PPP motifs and the active fragment (in total, residues 835–1026) binds to the PPP-binding site on Sec23a. Binding observed between Sec31 and wild-type Sec23a (lane 7) is substantially reduced by mutation of the PPP-binding site (lanes 8 and 9). (*B*) The reduction in binding to Sec31 caused by mutation to Sec23a was quantified by densitometry of bands in *A*. This figure can be usefully compared with Fig. 3*B*. (*C*) Binding of proteins to Sec23a was tested in the presence of Sar1 (human Sar1a, residues 25–198) complexed with the nonhydrolyzable GTP analog, GppNHp. Tight binding of Sec31 (residues 835–1026) to Sec23a requires Sar1-GppNHp (compare lanes 5 and 6). By contrast, TANGO1 (residues 1751–1907) binding to Sec23a is independent of Sar1 (lanes 7 and 8). (*D*) Surface representation of a COPII complex in which we model the bivalent interaction of Sec31 to Sec23a, as indicated by the results of binding analysis. Residues $^{945}$PPP$^{947}$ of human Sec31a are bound to the gelsolin-like domain of Sec23a, followed by a 32-residue linker (dotted line) and the active fragment (residues 980–1015). This is a composite model of Sec23a/24d/PPP taken from the present analysis, plus Sar1-GppNHp and the active fragment of Sec31 (from the yeast crystal structure of Sec23•Sar1 complexed with Sec31 active fragment; PDB ID code 2QTV). Coloring is the same as Fig. 2*C*. (*E*) Model showing how the TANGO1/cTAGE5 protein might promote the growth of the Sec23/24•Sar1 helical lattice on a COPII tubule. The dimensions and symmetry of the COPII-coated tubule are replicated from the published tomography images in ref. 17. The locations of the repeat PPP motifs on the TANGO1 and cTAGE5 molecules are indicated by black dots. In actuality, TANGO1 has seven PPP motifs and cTAGE5 five PPP motifs, but the spacing of motifs suggests that both proteins could bind to four copies of Sec23/24•Sar1 in array. This assumes a general PPII helical character for the bound PRD regions, induced by the ~30% proline content.

bilayer surface (estimated from the tomography images presented in ref. 17).

The observation that TANGO1 and Sec31 both contain a PRD led to the suggestion that TANGO1 may mimic the binding mode of the Sec31 active fragment and, thereby, stall Sec13/31-catalyzed GTP hydrolysis on Sar1 (9). Although our data indicate interplay between the protein molecules (Fig. 4 *A*–*C*), the TANGO1 PPP motif clearly binds to a unique site on Sec23a. The idea that rapid GTP hydrolysis will mediate against assembly of large COPII coats is countered by the alternative model that Sec12 (ER-localized Sar1 nucleotide exchange factor) activity maintains sufficient steady-state levels of Sar1-GTP, and that the COPII coat remains stably bound to membrane regardless of dynamic Sar1 cycling (23, 24).

**A Consideration of COPII Lattice Adaptability.** Structural studies have established that the adaptability of the COPII coat (i.e., the capacity to adopt vesicular and tubular forms) depends on two molecular features: variable angles at the vertex, and centrally along the edge, of the Sec13/31 cage; and a flexible (unstructured) polypeptide connection between the inner and outer coat proteins (17, 22, 25). The variable angles enable the Sec13/31 assembly unit to adapt to cages of different curvature (in two dimensions), and the flexible link between Sec23/24•Sar1 and Sec13/31 allows for variation in the geometric relationship between the inner and outer coat proteins (17).

**Implications for the Role of TANGO1/cTAGE5 in COPII Coat Assembly.** Through these mechanisms the COPII coat is inherently adaptable, but what switches coat assembly from small spherical vesicles to large tubular carriers? The essential difference between the COPII coat on vesicles and tubules is the helical lattice of Sec23/24•Sar1 observed on the latter, which self-assembles to a close-packed structure, to yield a coat that contains supernumerary inner complexes (i.e., the stoichiometry of the coat layers is 2:1 of Sec23/24:Sec13/31, such that only half of all of the Sec23/24 proteins bind a Sec13/31 partner). The helical symmetry is incompatible with an isometric structure, thus the Sec23/24•Sar1 array is absent from COPII-coated vesicles, which have a more open and random arrangement of inner coat protein (17). In view of this finding, we propose that TANGO1/cTAGE5 stimulates the growth of COPII-coated tubules by promoting the addition of assembly units to the helical Sec23/24•Sar1 array (Fig. 4*E*). The repeat distance between gelsolin-like domains on a COPII-coated tubule (78 Å along the helical lattice line) (see, for example, ref. 17) and the distances between PPP motifs in the PRDs would allow TANGO1 and cTAGE5 to bridge as many as four copies of Sec23/24•Sar1 (Fig. 4*E*). Although TANGO1/cTAGE5 PRD sequences are rather divergent across species, the high proline content and the PPP motifs are present throughout (note that we cannot rule out the possibility that XPP or PPX are binding motifs, where X is a residue that favors the PPII helical conformation, such as glycine or alanine). According to our model, the arrangement of multiple weak-binding motifs spread

across the inherently flexible PRDs will allow the two arms of TANGO1/cTAGE5 to be mobile and easily displaced from individual binding sites on Sec23/24•Sar1, enabling the receptor to adopt multiple binding modes and alternately stabilize the Sec23/24•Sar1 lattice and capture additional Sec23/24•Sar1 assembly units (Fig. 4*E*).

The model that we have proposed has significant parallel with the mechanism of action of tau protein on the microtubule helical lattice. Four repeated binding motifs in tau are separated by flexible linker sequences, and the binding of tau to multiple tubulin units stabilizes and enhances microtubule lattice polymerization without impeding microtubule dynamics (26, 27).

**Stepwise COPII Coat Assembly and Compartmentation of TANGO1/cTAGE5 to the ER Membrane.** Puzzlingly, TANGO1 does not depart with cargo in the COPII-coated carrier (9). The finding that PPP motifs are present in Sec31 suggests that the recruitment of the Sec13/31 outer coat could displace TANGO1/cTAGE5 and confine it the ER membrane at the base of the growing tubule. COPII-coat assembly on membranes occurs stepwise—Sar1-GTP recruits Sec23/24, which recruits Sec13/31—because the binding site for Sec31 comprises a composite surface of the Sec23 and Sar1-GTP molecules (1, 22) (Fig. 4*C*, lanes 5 and 6). Thus, the compartmentation of TANGO1/cTAGE5 to the base of the growing tubule may be a spatial consequence of the stepwise assembly process.

Finally, the model we have proposed for TANGO1/cTAGE5 function envisages two self-assembly reactions competing to influence the shape of the COPII carrier: Sec23/24•Sar1 lattice formation to propagate the tubule, and Sec13/31 lattice formation to impose an isometric cage or to terminate a tubule. Accordingly, TANGO1/cTAGE5 would bias coat assembly toward tubulation by increasing the local concentration of Sec23/24•Sar1 assembly units, and stabilizing the lattice, at the base of the tubule (Fig. 4*E*). Likewise, other protein factors could influence the shape of the forming COPII carrier by regulating the availability of inner or outer COPII proteins at ER exit sites. Candidates for such a function include the Sedlin component of the TRAPP complex (28), the ubiquitin ligase CUL3-KLHL12 (29), and $Ca^{2+}$/ALG-2 (30, 31), which binds to a site on Sec31 PRD that includes one of the two PPP motifs, residues $^{838}$PPP$^{841}$ (32). The biochemical analysis and the mechanism that we have presented for TANGO1/cTAGE5 interplay with COPII should provide a basis on which to further explore the role of these protein factors in ER-to-Golgi transport.

## Materials and Methods

**Protein and Production and X-Ray Crystallography.** COPII coat proteins were expressed in baculovirus-infected insect cells. Sar1, TANGO1, and Sec31 polypeptides were overproduced in *Escherichia coli*. Proteins were purified by affinity and size-exclusion chromatography. Crystals of Sec23a bound to TANGO1 (residues 1780–1840) were obtained by vapor diffusion using a crystallization solution comprising 100 mM Hepes pH 7.5, 6% (vol/vol) isopropanol, and 300 mM sodium citrate. The crystal structure was determined by molecular replacement and refined to 2.6 Å resolution. Details are provided in *SI Materials and Methods*.

**Other Methods.** Binding experiments are described in *SI Materials and Methods*.

**ACKNOWLEDGMENTS.** We thank staff of the Northeast Collaborative Access Team beamlines at the Advanced Photon Source of the Argonne National Laboratory for access to synchrotron facilities. This research was supported in part by the Memorial Sloan Kettering Cancer Center Core Grant P30-CA008748.

1. Matsuoka K, et al. (1998) COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell* 93(2):263–275.
2. Barlowe C, et al. (1994) COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77(6):895–907.
3. Stagg SM, et al. (2006) Structure of the Sec13/31 COPII coat cage. *Nature* 439(7073): 234–238.
4. Zeuschner D, et al. (2006) Immuno-electron tomography of ER exit sites reveals the existence of free COPII-coated transport carriers. *Nat Cell Biol* 8(4):377–383.
5. Lee MC, Miller EA, Goldberg J, Orci L, Schekman R (2004) Bi-directional protein transport between the ER and Golgi. *Annu Rev Cell Dev Biol* 20:87–123.
6. Boyadjiev SA, et al. (2006) Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nat Genet* 38(10):1192–1197.
7. Garbes L, et al. (2015) Mutations in SEC24D, encoding a component of the COPII machinery, cause a syndromic form of osteogenesis imperfecta. *Am J Hum Genet* 96(3):432–439.
8. Townley AK, et al. (2008) Efficient coupling of Sec23-Sec24 to Sec13-Sec31 drives COPII-dependent collagen secretion and is essential for normal craniofacial development. *J Cell Sci* 121(Pt 18):3025–3034.
9. Saito K, et al. (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell* 136(5):891–902.
10. Saito K, et al. (2011) cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites. *Mol Biol Cell* 22(13):2301–2308.
11. Wilson DG, et al. (2011) Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse. *J Cell Biol* 193(5):935–951.
12. Nogueira C, et al. (2014) SLY1 and Syntaxin 18 specify a distinct pathway for procollagen VII export from the endoplasmic reticulum. *eLife* 3:e02784.
13. Santos AJ, Raote I, Scarpa M, Brouwers N, Malhotra B (2015) TANGO1 recruites ERGIC membranes to the endoplasmic reticulum for procollagen export. *eLife* 4:e10982.
14. Bannykh SI, Rowe T, Balch WE (1996) The organization of endoplasmic reticulum export complexes. *J Cell Biol* 135(1):19–35.
15. Watson P, Stephens DJ (2005) ER-to-Golgi transport: Form and formation of vesicular and tubular carriers. *Biochim Biophys Acta* 1744(3):304–315.
16. Bacia K, et al. (2011) Multibudded tubules formed by COPII on artificial liposomes. *Sci Rep* 1:17.
17. Zanetti G, et al. (2013) The structure of the COPII transport-vesicle coat assembled on membranes. *eLife* 2:e00951.
18. Bi X, Corpina RA, Goldberg J (2002) Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature* 419(6904):271–277.
19. Peterson FC, Volkman BF (2009) Diversity of polyproline recognition by EVH1 domains. *Front Biosci (Landmark Ed)* 14:833–846.
20. Zarrinpar A, Lim WA (2000) Converging on proline: The mechanism of WW domain peptide recognition. *Nat Struct Biol* 7(8):611–613.
21. Mancias JD, Goldberg J (2008) Structural basis of cargo membrane protein discrimination by the human COPII coat machinery. *EMBO J* 27(21):2918–2928.
22. Bi X, Mancias JD, Goldberg J (2007) Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. *Dev Cell* 13(5): 635–645.
23. Futai E, Hamamoto S, Orci L, Schekman R (2004) GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on synthetic liposomes. *EMBO J* 23(21):4146–4155.
24. Sato K, Nakano A (2005) Dissection of COPII subunit-cargo assembly and disassembly kinetics during Sar1p-GTP hydrolysis. *Nat Struct Mol Biol* 12(2):167–174.
25. Noble AJ, et al. (2013) A pseudoatomic model of the COPII cage obtained from cryo-electron microscopy and mass spectrometry. *Nat Struct Mol Biol* 20(2):167–173.
26. Butner KA, Kirschner MW (1991) Tau protein binds to microtubules through a flexible array of distributed weak sites. *J Cell Biol* 115(3):717–730.
27. Drechsel DN, Hyman AA, Cobb MH, Kirschner MW (1992) Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau. *Mol Biol Cell* 3(10):1141–1154.
28. Venditti R, et al. (2012) Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science* 337(6102):1668–1672.
29. Jin L, et al. (2012) Ubiquitin-dependent regulation of COPII coat size and function. *Nature* 482(7386):495–500.
30. Helm JR, et al. (2014) Apoptosis-linked gene-2 (ALG-2)/Sec31 interactions regulate endoplasmic reticulum (ER)-to-Golgi transport: A potential effector pathway for luminal calcium. *J Biol Chem* 289(34):23609–23628.
31. la Cour JM, Schindler AJ, Berchtold MW, Schekman R (2013) ALG-2 attenuates COPII budding in vitro and stabilizes the Sec23/Sec31A complex. *PLoS One* 8(9):e75309.
32. Takahashi T, et al. (2015) Structural analysis of the complex between penta-EF-hand ALG-2 protein and Sec31A peptide reveals a novel target recognition mechanism of ALG-2. *Int J Mol Sci* 16(2):3677–3699.
33. Otwinowski W, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276(Part A):307–326.
34. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.