Cell, Vol. 114, 395–399, August 22, 2003, Copyright ©2003 by Cell Press

Previews

Molecular Recognition of Cargo by the COPII Complex: A Most Accommodating Coat

The molecular mechanism by which diverse cargo proteins are recognized and exported from the ER has been unclear. Two papers in this issue of *Cell* (Mossessova et al., 2003; Miller et al., 2003) add clarity by mapping multiple cargo recognition sites in the Sec24 subunit of the COPII coat complex and demonstrating roles for these sites in export of specific protein cargos from the ER.

The coat protein complex II (COPII) catalyzes transport vesicle formation from the ER and segregates export cargo from ER-resident proteins. During the ER export stage, a remarkable variety of cargo molecules, including secreted growth factors, cell surface receptors, membrane channels, and extracellular matrix proteins, must be accommodated by the COPII machinery. And while some exit signals have been identified in exported proteins, such as the di-acidic or DXE motif (Nishimura and Balch, 1997), a single exit signal has not been defined that could explain efficient export of all known cargos. Indeed, a number of studies have suggested that the COPII-dependent export machinery can decipher several different signals. How might the COPII coat discriminate among the assortment of ER proteins for inclusion of specific cargos into transport vesicles? In a tour de force that combines crystallography, biochemistry, yeast genetics, morphology, and reconstituted in vitro transport assays, the Goldberg and Schekman labs now provide a molecular view of cargo recognition by the COPII coat.

COPII coats are assembled from three components: Sar1, Sec23/24 complex, and Sec13/31 complex. Activation of the Sar1 GTPase initiates coat assembly followed by sequential recruitment of Sec23/24 and then Sec13/31. Current models for protein sorting during export from the ER propose that prebudding complexes consisting of Sec23/24-Sar1 bound to export cargo form at ER exit sites and are then gathered into a polymerized COPII structure by the extended arrangement of the Sec13/31 complex (Antonny and Schekman, 2001). Several lines of evidence have implicated the Sec24 subunit of this coat complex in cargo recognition, acting as a putative cargo adaptor protein (Miller et al., 2002). The Sec23 subunit serves a regulatory role as a Sar1 specific GTPase-activating protein (GAP) and thus the Sec23/24 complex appears to link the activities of cargo selection and regulated coat assembly. Recent crystallographic analysis of a prebudding complex consisting of Sec23/24 bound to Sar1 revealed a "bowtie-shaped" structure with a concave structure proposed to conform to a curved membrane surface (Bi et al., 2002). These initial structural studies suggested the presence of potential cargo binding sites on the concave surface of this complex; however, vast regions of the Sec23/24 surface were uncharacterized.

To locate specific cargo interaction sites in the prebudding complex, Mossessova et al. (2003) first devised a Sec23/24 binding assay to identify binding sequences contained within some well established vesicle cargo proteins, Sed5 and Bet1p. These ER/Golgi SNARE proteins are efficiently packaged by the COPII coat and are needed for subsequent membrane fusion stages. They identified and refined the SNARE binding sequences then designed synthetic peptides that corresponded to the Sed5 or Bet1 binding signals. Co-crystals of SNARE peptides bound to Sec24 were then analyzed by X-ray crystallography, allowing the investigators to map electron density contributed by the peptides onto the Sec24 structure. Strikingly, they observed two independent binding sites in Sec24 that recognize distinct SNARE peptide sequences. These two sites are separated by about 80 Å and were described as the "A-site" and "B-site" (see Figure 1). The A-site has specificity for a signal in Sed5 whereas the B-site binds to a distinct set of export signals contained in Bet1 and a second lower affinity signal in Sed5. Furthermore, binding of a known di-acidic export signal overlapped with the B-site and structural analyses showed that binding depended on some of the same contacts as the Bet1 sequence. However, the Sec24 contacts between residues in the di-acidic sequence compared to the Bet1 sequence were notably distinct, indicating that the manner in which related signals bind to the same cargo recognition site can be flexible. Finally, the authors find that even when both the A- and B-sites are occupied, a third ER/Golgi SNARE protein, Sec22, binds to Sec24, indicating the presence of additional cargo binding sites.

Mossessova and colleagues also explored the interactions between Sec24 and ER/Golgi SNARE proteins in various assembled and semi-assembled states. Cognate sets of SNARE proteins are known to form stable complexes through assembly of their SNARE motifs into a parallel four-helix coiled-coil structure. Assembly of SNARE complexes in *trans* (i.e., between donor and acceptor membranes) mediates fusion of intracellular membranes (Jahn et al., 2003). The Sec24 binding experiments in the current study indicated that the Bet1 binding signal was occluded when assembled into a SNARE complex with Sed5, Sec22, and Bos1. In contrast, a high affinity COPII binding signal in Sed5 was available only when assembled into SNARE complexes and apparently sequestered in a closed conformational state of monomeric Sed5. Based on these and other experimental results, the authors postulate that the COPII coat selects certain combinations of SNARE proteins to program vesicles for future membrane fusion events.

In the companion report, Miller et al. (2003) investigate the consequences of mutating specific Sec24 resides on COPII-dependent cargo selection and on cell viability. In this study, the authors initially performed an alanine scan of conserved residues based on the Sec23/24 crystal structure (Bi et al., 2002). Interestingly, they encoun-

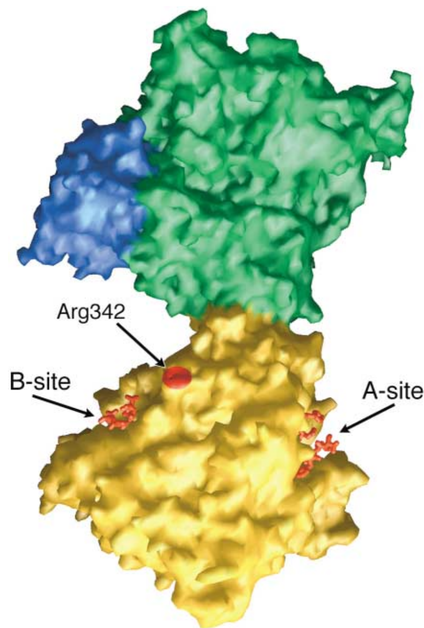

Figure 1. Multiple Cargo Binding Sites in the Sec23/24-Sar1 Prebudding Complex

Surface representation of the prebudding complex with the Sec23 subunit colored in green, Sec24 in yellow, and Sar1 in blue. Red highlights in Sec24 indicate positions of the YNNSNPF peptide sequence of Sed5 bound at the A-site, the LASLE peptide sequence of Bet1 bound at the B-site and arginine residue 342 (Arg342) that is required for Sec22 recognition. A characterized di-acidic export signal (DLESQ) also binds to the B-site of Sec24. (Figure kindly provided by J. Goldberg).

tered several of the same amino acid residues implicated in cargo binding at the B-site by Mossessova and colleagues. Expanding on this initial analysis, the investigators generated a collection of mutant Sec24 proteins that included other residues lining the B-site binding pocket. They purified this set of mutant proteins to test activity in established in vitro cargo binding and COPII budding assays. All of the B-site mutants efficiently packaged the precursor form of α-factor, a soluble secretory protein, into COPII vesicles at wild-type levels. These results indicate the Sec24 mutations do not interfere with COPII coat assembly or membrane budding. However, other vesicle proteins were inefficiently sorted into COPII vesicles when B-site mutants were used to drive budding reactions. Notably, Bet1 and other ER/Golgi SNARE proteins were inefficiently packaged into vesicles, resulting in a reduced fusion competence of vesicles with Golgi acceptor membranes. These findings nicely corroborate the structural studies and demonstrate a physiological relevance of the B-site in cargo recognition.

Expression of specific Sec24 B-site mutants as the sole source of Sec24 in yeast cells produced a range of phenotypes from inviability and temperature sensitivity to apparent normal growth. However, many of the Sec24 mutant phenotypes were exacerbated when combined with deletion of *ISS1* and/or *LST1*, two nonessential genes that encode Sec24-related proteins in *S. cerevisiae*. These homologs are required for export of certain secretory cargo from the ER and display both distinct and overlapping functional properties with Sec24 (Miller et al., 2002). Aware of the synthetic growth phenotypes of the Sec24 B-site mutants when combined with *iss1*Δ and/or *lst1*Δ, the investigators examined other alanine scanned Sec24 mutants in this background. One of these mutations, in arginine 342, was lethal when combined with the *iss1*Δ allele. Further characterization of this mutant Sec23/24 complex revealed a distinct sorting defect in reconstituted COPII budding assays. The arginine 342 mutation produced a 90% reduction in packaging of Sec22 into budded COPII vesicles. The defect was quite specific as other ER/Golgi SNAREs and all other cargo proteins monitored were incorporated into vesicles at normal levels. Based on these observations, the authors conclude that arginine 342 resides in a third distinct cargo recognition site in Sec24 (Figure 1) in accord with the Mossessova et al. results on Sec22 binding.

In summary, these reports identify multiple cargo binding sites in the Sec24 protein and define molecular contacts between specific export signals and this cargo adaptor. Do we now understand the mechanism of cargo export from the ER? Well, not entirely. As with many processes in biology, binding is only a part of the mechanism and must be reversed in subsequent steps. In the case of ER export, one of the key regulatory factors, Sar1, has not been fully placed into the scheme. Cargo binding is presumably regulated in a manner such that Sec24 binds to ER localized export cargo but is released during vesicle uncoating and/or fusion stages. Here, the Sar1 GTPase probably plays a critical role either in contacting cargo directly or by attracting coat subunits to the membrane surface. The current reports and others have pointed out that certain cargo proteins contain multiple ER export signals that may be recognized by Sar1 or by other regions of the Sec23/24 complex (Otte and Barlowe, 2002). A combination of export signals may be needed for efficient entry of specific assemblies of proteins into COPII vesicles as proposed by Mossessova and colleagues for SNARE complexes and may also provide an important regulatory mechanism for control of cellular homeostasis (Yang et al., 2002). Further studies will be required to address these regulatory issues. Regardless, an abundance of Sec24 cargo recognition sites described in these reports combined with an extended family of Sec24 proteins, begins to account for the spectrum of protein cargo that must be efficiently exported from the ER.

Charles Barlowe
Dartmouth Medical School
Department of Biochemistry
Hanover, New Hampshire 03755

Selected Reading

Antonny, B., and Schekman, R. (2001). Curr. Opin. Cell Biol. *13*, 438–443.

Bi, X., Corpina, R.A., and Goldberg, J. (2002). Nature *419*, 271–277.

Jahn, R., Lang, T., and Sudhof, T.C. (2003). Cell *112*, 519–533.

Miller, E., Antonny, B., Hamamoto, S., and Schekman, R. (2002). EMBO J. *21*, 6105–6113.

Miller, E., Beilharz, T., Malkus, P., Lee, M., Hamamoto, S., Orci, L., and Schekman, R. (2003). Cell *114*, this issue, 497–509.

Mossessova, E., Bickford, L.C., and Goldberg, J. (2003). Cell *114*, this issue, 483–495.

Nishimura, N., and Balch, W.E. (1997). Science *277*, 556–558.

Otte, S., and Barlowe, C. (2002). EMBO J. *21*, 6095–6104.

Yang, T., Espenshade, P.J., Wright, M.E., Yabe, D., Gong, Y., Aebersold, R., Goldstein, J.L., and Brown, M.S. (2002). Cell *110*, 489–500.

Sex Determination in the Honeybee

Sex determination in honeybees involves a multi-allelic locus, such that homozygotes develop as males and heterozygotes as females. In this issue of *Cell*, Beye and colleagues (2003) report the cloning of the sex-determining gene, *csd*. It codes for an SR protein, and different alleles have very different amino-acid sequences. Inactivating *csd* leads to development as a male.

In 1845, the German apiarist J. Dzierzon proposed that male honeybees arise from unfertilized eggs, while females come from fertilized eggs. It is now known that this sex determination system (haplodiploidy) is probably common to all sexually reproducing members of the *Hymenoptera* (ants, bees, and wasps) and *Thysanoptera* (thrips), as well as being found sporadically in other orders of insects including beetles and *Homoptera* (scale insects and whiteflies), in the *Acarina* (ticks and mites), and in monogont rotifers (Bull, 1983). It is very hard to see how haplodiploidy could evolve from one of the two best-known sex-determining systems, a male-determining Y chromosome (as found in mammals), or X: autosome balance (as in *Drosophila* and *Caenorhabditis*). In the first case, it is impossible for a male to develop without the male-determining region of the Y, and for a female to develop in its presence. In the second case, there is no difference in X: autosome balance between haploids and diploids.

In the few cases (all in *Hymenoptera*) in which a detailed genetic analysis of haplodiploid sex determination has been performed, the mechanism involves what is known as complementary sex determination (Bull, 1983). Females are always heterozygous for a pair of distinct alleles at the sex-determining locus, whereas males are homozygous (if derived from a fertilized egg), or haploid (if derived from an unfertilized egg). Such a sex determination system can perfectly well exist without haplodiploidy if all eggs are fertilized, and evolutionary models of the conversion of diploid complementary sex determination into haplodiploidy can be constructed (Bull, 1983). This removes some of the mystery surrounding the origin of haplodiploidy.

In honeybees, the best-studied example of complementary sex determination, homozygous diploid males can be produced by inbreeeding, but are normally eaten by the workers. The lethality of diploid males means that there is a selection pressure to increase the number of functionally distinct alleles, and indeed in honeybees as many 12 alleles have been detected in a single population (Bull, 1983). This pressure to make the system highly polymorphic is similar that in the self-incompatibility loci of flowering plants and some fungal mating types, where a large number of alleles coexist and successful matings only occur between individuals carrying different alleles (Charlesworth, 2002; Casselton, 2002). Variability with respect to neutral markers at sites closely linked to the sex locus itself has been exploited by Beye et al. to identify, clone, and characterize the sex-determining gene of honeybees, *csd*, in a tour de force of positional cloning (Beye et al., 2003).

Earlier work had identified two genetic markers flanking the sex-determining locus, at distances of 1 and 7 cM. Using the closer marker, they isolated 70 kb of DNA containing the sex locus, using polymorphic markers to orientate a chromosome walk. They found a 13 kb region that was always heterozygous in females in the cross they used. cDNA analyses identified a transcript within this region, and they inferred that this was likely to be the sex-determining gene itself. Sequencing of the corresponding genomic DNA shows that *csd* consists of 1453 bases, with nine exons and a protein of 385 amino acids. The protein is a novel type of arginine-serine rich (SR) protein. Intriguingly, its C terminus has some sequence similarity with the protein coded by the *tra* locus of *Drosophila*, an important player in sex determination (Marín and Baker, 1998). Expression studies showed that *csd* is transcribed in both males and females, starting at 12 hr of development, so that differential expression plays no role in sex determination.

Sequence comparisons were carried out between four different sex-determining alleles, revealing an unusually high level of amino acid sequence differences between alleles, especially toward the C-terminal region. In addition to single substitutions, alleles typically differ with respect to insertions and deletions of amino acids, often in tracts of several at a time. Variation in a hypervariable region mainly involves repeats of asparagine and tyrosine residues. This abundance of amino acid sequence differences among alleles is also seen in self-incompatibility alleles, and presents a considerable problem for identifying sites of functional significance. The magnitude of these differences suggests that the alleles may be rather old, in terms of evolutionary origin; it would be of interest to look for trans-specific polymorphisms of the type sometimes found at self-incompatibility loci (Charlesworth, 2002).

This pattern of variation in itself strongly supports the inference that *csd* is the sex-determining locus. Further evidence is provided by functional studies, using RNA interference. Injection of *csd* dsRNA into developing eggs caused genetic females to develop as male larvae with high probability, but males were unaffected. This indicates that *csd* function is required in females, but not in males. In turn, this implies that its product is nonfunctional when *csd* is transcribed from a single allele. Again, this has some parallels with self-incompatibility and mating-type systems, in which pollination or development of a sexual fusion product cannot proceed if genetically similar partner cells are involved. The fact