# THE ENZYMES

## Molecular Machines Involved in Protein Transport across Cellular Membranes

Edited by
Ross E. Dalbey
Carla M. Koehler
Fuyuhiko Tamanoi

VOLUME XXV

# *The Enzymes*

VOLUME XXV

## *MOLECULAR MACHINES INVOLVED IN PROTEIN TRANSPORT ACROSS CELLULAR MEMBRANES*

This page intentionally left blank

# THE ENZYMES

Edited by

Ross E. Dalbey

*Department of Chemistry*
*The Ohio State University*
*Columbus, OH 43210, USA*

Carla M. Koehler

*Department of Chemistry and Biochemistry*
*University of California, Los Angeles*
*Los Angeles, CA 90095, USA*

Fuyuhiko Tamanoi

*Department of Microbiology, Immunology, and Molecular Genetics*
*Molecular Biology Institute*
*University of California, Los Angeles*
*Los Angeles, CA 90095, USA*

Volume XXV

## *MOLECULAR MACHINES INVOLVED IN PROTEIN TRANSPORT ACROSS CELLULAR MEMBRANES*

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
84 Theobald's Road, London WC1X 8RR, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2007



ISBN: 978-0-12-373916-2
ISSN: 1874-6047



Printed and bound in USA

07 08 09 10 11 10 9 8 7 6 5 4 3 2 1

# *Contents*

## 4. YidC: A Protein with Multiple Functions in Bacterial Membrane Biogenesis

NIL CELEBI AND ROSS E. DALBEY



## 5. Disulfide Bond Formation Enzymes

JAMES C.A. BARDWELL



## 6. The Identification of the YaeT Complex and Its Role in the Assembly of Bacterial Outer Membrane β-Barrel Proteins

JULIANA MALINVERNI AND THOMAS J. SILHAVY

## Part IV: Crossing Chloroplast Membranes

### 16. The Toc Machinery of the Protein Import Apparatus of Chloroplasts

Caleb Rounds, Fei Wang, and Danny J. Schnell



### 17. The Role of the Tic Machinery in Chloroplast Protein Import

Philipp Benz, Jürgen Soll, and Bettina Bölter



### 18. The Sec and Tat Protein Translocation Pathways in Chloroplasts

Kenneth Cline and Steven M. Theg

This page intentionally left blank

# *Preface*

Proteins are synthesized in the cytosol but need to be transported to proper cellular locations for their action. How this intracellular transport is accomplished is one of the central issues in biology. We now know that this process involves targeting sequences on the protein as well as machines (multiprotein complexes) in the membrane that facilitate transport of proteins across cellular membranes. Recently, we have seen dramatic advances in our understanding of various machines involved in protein transport across cellular membranes in prokaryotes and eukaryotes. This volume is intended to bring together a diverse range of study involving molecular machines for protein transport across membranes. While previous publications dealt more with targeting signals on the protein or biogenesis of organelles, we tried to focus our attention on the machines that operate to facilitate uptake of proteins across the membrane. Our hope is that the volume will convey wealth of knowledge on diverse machines that function in various cellular membranes to facilitate protein transport. Similarities and differences of these systems are discussed.

Chapters are grouped into five different sections. Part I focuses on bacterial membranes. First, machines involved in the membrane targeting and transport through the inner membrane such as SRP, Sec, YidC, and Tat are discussed. This is followed by the discussion on the machines in the outer membrane. Protein transport into endoplasmic reticulum is discussed in Part II. Topics in this section cover SRP and Sec as well as chaperones such as BiP/ker2p and ERp57. Part III deals with the transport of proteins into mitochondria and describes inner membrane and outer membrane machines. TOM and SAM complexes as well as TIM complex are discussed. Part IV describes protein transport into chloroplast in plants. TOC and TIC complexes operate at the outer and inner envelope of chloroplast, respectively. In addition, Sec and TAT pathways function to transport proteins across the thylakoid membrane into the lumen. In the final section (Part V), we discuss mechanisms of peroxisomal protein import.

Because unfolding is one of the important features of protein transport process, two chapters (Chapters 5 and 13) discuss the significance of disulfide bond formation. In addition, some chapters include discussions on the comparison of the Sec pathway that transports substrates in an unfolded

state and the Tat pathway that transports substrates in a folded state. It is also important to point out that similar transport mechanisms operate in different systems. For example, SRP and Sec complexes are used in bacterial systems as well as in the import into endoplasmic reticulum of eukaryotic systems. The Tat machinery is present in bacterial systems as well as in chloroplasts in plants. Chapters 3 and 18 contain discussions on this conservation of transport machines and discussion on how studies in bacterial and plant systems facilitated our understanding of this transport machine.

The idea for this volume was conceived by one of us (R.D.). We have been able to bring this initial idea to publication in a short period of time. This is in large part due to the effort of the contributors to prepare their chapters in a timely fashion. We would like to thank the authors for making this publication a reality. We also thank Tari Broderick and Renske van Dijk for their expert help in the preparation of this volume.

Fuyuhiko Tamanoi
Ross Dalbey
Carla Koehler
April 20, 2007

# Part I

## *Crossing Bacterial Membranes*

This page intentionally left blank

# 1

# *Cotranslational Protein Targeting in* Escherichia coli

RONALD S. ULLERS[a] • PIERRE GENEVAUX[b] • JOEN LUIRINK[c]

[a]*Department of Microbiology and Molecular Medicine*
*Centre Médical Universitaire*
*CH-1211 Geneva, Switzerland*

[b]*Laboratoire de Microbiologie et Génétique Moléculaires*
*IBCG, CNRS, Université Paul-Sabatier*
*118, route de Narbonne, 31062 Toulouse cedex 09, France*

[c]*Department of Molecular Microbiology*
*Institute of Molecular Cell Biology*
*Vrije Universiteit*
*1081 HV Amsterdam, The Netherlands*

## I. Introduction

In all organisms, genetic messages are translated primarily by cytosolic ribosomes, yet the translation products end up in a variety of cellular locations. Nascent polypeptide chains emerging from the ribosomal exit are facing a pool of chaperones, folding catalysts, and targeting factors, which efficiently partition between proteins that need to be folded in the cytosol and proteins that need to be transferred into or through membranes. In Gram-negative bacteria, proteins destined for the secretory apparatus are equipped with an N-terminal extension called the signal peptide (reviewed in [1]). The signal peptide directs the secretory proteins (periplasmic and outer membrane proteins) into the SecB pathway, where, during or after translocation through the inner membrane (IM), the signal



DOI: 10.1016/S1874-6047(07)25001-2

peptides are cleaved off by signal peptidases. Most inner membrane proteins (IMPs) lack a cleavable signal peptide. They are anchored into the IM by hydrophobic α-helical transmembrane (TM) segments. These TM segments may also act as targeting signals in which case they are recognized by the signal recognition particle (SRP) that directs the polypeptides to the Sec complex at the IM (reviewed in [2, 3]). In contrast to translocation of unfolded proteins across the IM via the Sec system, the twin-arginine translocation (TAT) system can only transport proteins that are completely folded (reviewed in [4]).

The decision on a protein's destiny is probably made early during its synthesis on the ribosome, since both SRP and chaperones such as trigger factor (TF) bind close to the ribosomal exit site where they can associate with the emerging nascent chain. In this chapter, we will focus on the different stages of SRP-mediated protein targeting in bacteria—from the synthesis on the ribosome to the final handover of the protein to the SecYEG complex in the IM—and on the complexity of interactions between the newly synthesized polypeptides and cellular factors that control their destiny.

## II. The Ribosome

Protein synthesis is catalyzed by the ribosome, a highly conserved macromolecular complex present in all living cells [5]. Both the small and large subunit of the ribosome is composed of RNA and proteins. The small subunit mainly decodes genetic information [6, 7], while the large subunit is responsible for peptide elongation and protein release. Peptide bond formation occurs in the peptidyl transferase center (PTC) [8, 9].

### A. The Ribosomal Tunnel

Crystal structures of archaeal and eubacterial large ribosomal subunits [8–12] show a long tunnel running from the PTC to the ribosomal proteins L23/L24/L29 at the surface of the ribosome. It has been suggested that this is the normal exit path from the ribosome for nascent peptides [10]. However, some polypeptides may leave the PTC via the interface of the two ribosomal subunits [13]. In addition, three-dimensional cryo-electron microscopy (EM) maps of the ribosome reveal several side branches of the tunnel that may reflect alternative exit sites [14, 15]. Nevertheless, cross-linking studies suggest that nascent chains, irrespective of their future destination, exit the ribosome near L23 [16–19].

The length of the main tunnel is about 100 Å, and its diameter varies from 10 to 20 Å. The inner surface of the tunnel consists mainly of rRNA, but nonglobular parts of the ribosomal proteins L4 and L22 also contribute [9, 10]. The resulting surface is largely hydrophilic and uncharged, thus facilitating the passage of all kinds of peptide sequences [10]. A long β-hairpin loop of L22 lies approximately parallel to the tunnel axis, making this ribosomal protein the largest protein contributor to the inner surface of the tunnel. The ribosomal proteins L23, L24, and L29 flank the exit of the tunnel.

The tunnel was initially thought to be a narrow passage through a rigid structure that precludes significant protein folding [10]. Structural studies, however, suggest that the ribosome is rather dynamic and adopts different functional conformations [15, 20–22]. In addition, several reports have demonstrated that nascent proteins fold to various degrees inside the ribosome [23–26]. Consequently, the ribosomal tunnel must expand considerably during protein synthesis to accommodate folded nascent polypeptides [24, 27]. However, this view has been questioned by Voss *et al*. [28]. On the basis of geometric analysis of the ribosomal tunnel in *Haloarcula marismortui*, it was proposed that the tunnel is not wide enough to accommodate folded polypeptides larger than α-helices [28].

Although the ribosomal tunnel wall has a ''nonstick'' character [10], certain nascent chains can interact with the tunnel, thereby causing translational pausing (reviewed in [13, 29, 30]). One striking example is the SecM protein [31]. SecM includes a 17 amino acid sequence motif that can block protein elongation and create a stalled ribosomal complex in the absence of a functional protein export system. It was demonstrated by fluorescence resonance energy transfer (FRET) that the SecM C-terminus adopts a compact conformation on synthesis of the arrest motif, which appeared essential for the translocation arrest and which was specifically induced by the ribosome [26]. It was proposed that translation arrest by SecM results from series of reciprocal interactions between the ribosome and the C-terminal part of the nascent SecM polypeptide. On the basis of cryo-EM reconstructions of pretranslocational and SecM-stalled *Escherichia coli* ribosome complexes, Mitra *et al.* [32] also suggested that SecM induces a cascade of ribosomal conformational changes that lock the mRNA-tRNA complex on the ribosome, such that elongation is stalled [32]. It was previously shown that SecM-mediated elongation arrest can be bypassed by mutations in the 23S rRNA or in L22 at a constriction area in the tunnel, which might act as a discriminating gate [31]. Taken together, the cryo-EM and the biochemical data suggest that nascent chain interactions and structural rearrangements in the ribosomal interior affect translation rates of nascent polypeptide chains.

### B. Sensing Nascent Secretory and Membrane Proteins in the Tunnel?

As discussed above, nascent peptides may have specific interactions already in the ribosome tunnel, which may regulate translation. In addition, specific interactions and conformational changes in the ribosome may be sensed and transduced to the surface of the ribosome influencing downstream interactions and topogenesis of the protein synthesized [13, 29, 30]. *In concreto*, future TM segments may be recognized already inside the ribosome. These intraribosomal contacts may have a profound impact on downstream processes such as the extraribosomal contacts with chaperones, targeting factors, and translocon components (reviewed in [33]). A fluorescence quenching study from the Johnson group has indicated that, in eukaryotes, a TM segment inside the ribosome induces conformational changes in the Sec translocon in the endoplasmic reticulum (ER) membrane [34]. How is the presence of a TM segment at the entrance of the ribosomal tunnel signaled to the exit site of the ribosome, and sequentially, causes a conformational change in the translocon?

This question has been addressed by using a combined FRET analysis and cross-linking approach [25]. It was found that nascent secretory proteins are fully extended inside the eukaryotic ribosome, whereas nascent membrane proteins (containing a TM segment) are folded into a more compact (presumably α-helical) structure [25]. Furthermore, it was demonstrated that nascent membrane proteins cross-linked to proteins L17 (the eukaryotic homologue of the *E. coli* L22) and L39 (not present in prokaryotes but partly replaced by L23) inside the ribosomal tunnel, while nascent secretory proteins did not [25]. It was proposed that the acquisition of α-helical structure by TM segments inside the ribosome is sensed by proteins L17 and L39 and signaled to the lumenal side of the Sec translocon, causing a closure of the translocon by the luminal chaperone Bip. In this way, the permeability barrier is maintained during membrane insertion of the nascent chain (reviewed in [35]). It has also been suggested that the Sec translocon itself is capable of maintaining the membrane barrier, being sealed either by a plugging domain of the closed translocon or by a translocating nascent chain inside the translocon (reviewed in [36]).

In *E. coli*, specific contacts of a TM segment in the ribosome have not been observed. L4, L22, and L23 are found cross-linked to nascent chains of the IMP Lep (irrespective of the presence of a functional TM segment; Figure 1.1), the secretory protein PhoE, and the cytoplasmic protein RpoB [17]. In contrast to the eukaryotic situation, in which both secretory and membrane proteins are cotranslationally targeted by the SRP to the Sec complex in the ER membrane, the targeting of secretory proteins in *E. coli* occurs mainly via a posttranslational SRP-independent process that

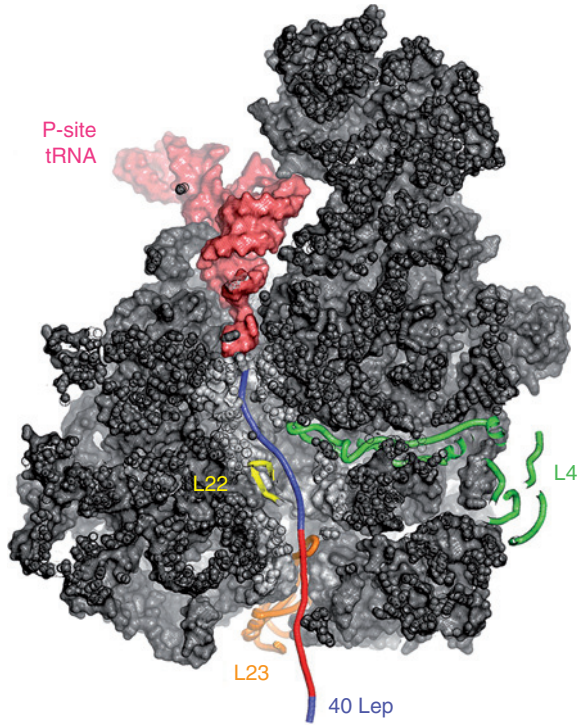



FIG. 1.1. The molecular environment of 40Lep in the ribosomal tunnel. The 40 amino acid residues nascent Leader Peptidase [Lep] (blue) and TM segment (red) is stretched from P-site tRNA (salmon pink) to the ribosomal exit site, and makes contact in the tunnel with ribosomal proteins L22 (yellow), L4 (green), and L23 (orange). (*Reprinted from* The Journal of Cell Biology *[17]*.) (See color plate section in the back of the book.)

involves the cytosolic chaperone SecB (Section V). Because cotranslational insertion at the *E. coli* Sec translocon seems to occur predominantly for IMPs, the interaction with the ribosome-nascent chain (RNC)–SRP complex might be sufficient to prime the Sec translocon for the membrane integration process. Alternatively, the translocon-associated protein YidC (not present in the ER), which specifically interacts with the TM segments of IMPs at an early stage of synthesis, can fulfil such a function. It should also be noted that a BiP homologue is not present in the *E. coli* periplasm and other plugging devices have not been described. Therefore, TM signaling in *E. coli* may not be critical at this stage.

Although the ribosomal signaling of an upcoming TM segment may be less relevant in *E. coli*, a recent cryo-EM analysis of eubacterial translocon-bound RNC complexes versus free translating and nontranslating

ribosomes showed that the ribosome adopts different conformations at the polypeptide exit on, first, recognition of a TM segment in a nascent FtsQ polypeptide, and second, binding to the translocon [37]. Positions of ribosomal elements—proteins L23/L29 and 23S rRNA hairpin h59—on one site of the exit site (forming one of the main connections with the translocon) remain fixed. In contrast, ribosomal elements on the other site of the exit site—forming another connection with the translocon and comprising proteins L4, L22, L24, and rRNA—display significant movement on recognition of a TM segment and binding to the translocon [37]. On the basis of these results, the authors propose a model in which conformational changes in the Sec translocon may be modulated by the ribosome through interactions of ribosomal RNA and proteins in the tunnel with a TM segment of a nascent peptide that will influence rRNA and protein dynamics at the ribosomal exit site [37].

### C. L23 at the Exit Site

The ribosomal protein L23, localized at the tunnel exit, plays a crucial role in the coordination of targeting and membrane insertion of IMPs. First, it is located adjacent to most, if not all, nascent peptides the moment they leave the ribosome [16–19, 38, 39]. Second, it serves as a docking protein for SRP [39, 40] and for TF [41, 42], an abundant cytosolic chaperone adjacent to most emerging nascent chains. Third, L23 is a major contributor to the junction between the ribosome and the Sec translocon during cotranslational membrane insertion in *E. coli*, yeast, and mammals [14, 43, 44].

## III. Chaperones and Targeting Factors at the Ribosomal Tunnel Exit

### A. SRP–Ribosome Interaction

SRP can bind to all ribosomes. It has significant affinity for both nontranslating ribosomes ($K_D$ ~80 nM in eukaryotes and prokaryotes) and ribosomes that carry a nascent chain either inside the ribosome ($K_D$ ~8 nM in eukaryotes) or outside the ribosomal tunnel exposing a targeting signal ($K_D$ ~0.2 nM in eukaryotes) [45–47]. The increase in binding affinity of SRP to ribosomes carrying a nascent chain inside the exit tunnel (before it emerges from the ribosome) over empty ribosomes [46] could provide a plausible mechanism to actively recruit SRP to the ribosomal peptide exit site. The recruitment of SRP to the ribosomal exit may help to increase the local concentration of SRP and might explain how SRP can function

efficiently, despite its low abundance in *E. coli* cells. In this way, SRP is allowed to scan and detect nascent chain targeting signals as soon as possible, thereby minimizing the chances that the polypeptide chain would become too long for efficiently targeting to the translocon [46, 48]. *In vitro* photo cross-linking experiments, however, suggest that SRP is not recruited to the ribosomal exit in *E. coli* in response to the presence of a TM segment inside the ribosomal tunnel. Nascent polypeptides carrying the first TM segment of the IMP Leader peptidase (Lep) in the exit tunnel did not induce SRP cross-linking to an arbitrary (nonhydrophobic) upstream sequence exposed outside the ribosome [17].

Cross-linking studies have revealed that L23 functions as an attachment site for the *E. coli* SRP on the ribosome [39, 40]. Likewise, the mammalian SRP54 could be cross-linked to L23a and L35, which are homologous to bacterial L23 and L29, respectively [49], and the yeast SRP54 could be coimmunoprecipitated specifically with L25, the homologue of bacterial L23 [50]. Combined, these data suggest that the ribosomal docking site for SRP is conserved. A three-dimensional cryo-EM map of the mammalian SRP bound to an active ribosome with a nascent secretory protein has been presented [51]. The structure reveals that the S-domain of SRP (containing the SRP core which is involved in signal peptide binding) contacts the large ribosomal subunit near the nascent chain exit site and that the Alu domain reaches into the elongation factor-binding site of the ribosome, in keeping with its elongation arrest activity. The structure suggests that the main connection between the S-domain and the ribosome is formed by SRP54 and L23a/L35. The tip of the SRP54 N-domain forms the contact with these ribosomal proteins, whereas the M-domain is located at the ribosomal exit site. This strategically positions SRP54 to scan emerging polypeptides for the presence of signal peptides (reviewed in [52]).

### B. TF–Ribosome Interaction

As mentioned above, the *E. coli* L23 protein functions as an attachment site not only for SRP but also for TF. TF is an abundant cytosolic protein of which a part is associated with the large ribosomal subunit in a 1:1 stoichiometry [53], and which possesses both chaperone and peptidyl-prolyl *cis*/*trans* isomerase (PPIase) activities [54–56]. As a *bona fide* chaperone, TF binds to unfolded substrate proteins, prevents their aggregation, and promotes their refolding [56–59]. The TF substrate-binding motif is composed of eight consecutive residues in which aromatic and basic residues are favored and acidic residues are disfavored [60]. On average, the TF-binding motifs appear every 32 amino acids in any protein and are mostly buried in the native protein structure [60]. Evidence has revealed that TF cooperates

in the folding of newly synthesized proteins with the downstream DnaK (Hsp70) chaperone, and that such cooperative functions were critical for bacterial survival [61–66]. As stated above, TF also displays a PPIase activity *in vitro* [55]. However, the presence of prolyl residues in both peptides or protein substrates does not significantly contribute to TF binding [60, 67], and the *in vivo* relevance of this activity remains to be determined [59, 65].

TF is the first chaperone to meet nascent polypeptides emerging from the ribosomal tunnel [16, 17, 19, 68–70]. TF was shown to cross-link to nascent polypeptides as short as 44 amino acid residues in length [17]. Binding of TF to substrates is critically dependent on their ribosomal attachment since the affinity for unfolded polypeptides in the absence of ribosomes is rather low ($K_D \sim 1\ \mu M$) [57, 67, 71]. TF binds with a similar affinity to vacant ribosomes ($K_D \sim 1\ \mu M$), but with slow kinetics compared to TF–polypeptide binding [53, 60, 72–74]. Consistent with previous data showing that complexes formed between TF and RNCs are highly salt-resistant [68], a study showed that the affinity of TF for ribosomes carrying nascent cytosolic substrates greatly increases, up to 20-fold, with increasing nascent chain length [74]. During translation, additional TF molecules are recruited to translating ribosomes, which appears to be dependent on the ability of TF to bind to the large ribosomal subunit [64]. The data suggest a ''cycling'' mechanism where TF interacts with the nascent chain and dissociates from the ribosome on growth of the nascent substrate. After release of TF from the ribosome, TF comigrates with the nascent chain, and new TF molecules bind to the same ribosome to interact with downstream substrate sequences [64].

*E. coli* TF is composed of three domains: the N-terminal domain (amino acids 1–144), the central PPIase domain (amino acids 145–247), and the C-terminal domain involved in TF chaperone function (amino acids 248–432) [75–77]. Ribosomal attachment is mostly mediated by a highly conserved GFRXGXXP motif, termed ''TF signature,'' in the N-terminal part of TF [41, 77]. The structure of *E. coli* TF bound to ribosomes has been modeled from the crystal structures of *E. coli* TF and the N-terminal fragment of *E. coli* TF bound to *Haloarcula* ribosomes [78]. The chimeric structure offers detailed insight into the exit portal. Noticeably, TF arches over the nascent chain exit site, forming an extended, hydrophobic cavity, which was proposed to shield nascent chains from aberrant intra- and intermolecular interactions [78]. In agreement with this model, Hoffmann *et al.* [79] have described the protection of rather large nonnative polypeptides from proteolytic cleavage by ribosome-bound TF in *E. coli* [79]. However, this protective effect appeared to depend on the nature of the nascent chain [80]. It should be noted that the composed structure might substantially differ from the genuine

situation in a number of ways. First, the TF–ribosome complex lacks a nascent chain, which may alter the orientation of TF domains with respect to each other and the ribosome. Second, protein L23 in eubacterial ribosomes exposes an elongated loop, which is not present in archaeal L23, that penetrates from the subunit exterior into the tunnel wall, and which may affect the interaction with TF [22, 81]. Third, protein L24, which is shorter in archaeal ribosomes, contacts TF on the *E. coli* ribosome, and may diminish the proposed molecular cradle of TF [82]. Clearly, more structural data from homologous systems are needed to clarify the position of TF on the large ribosomal subunit, and the structural rearrangements it might undergo on the synthesis of a polypeptide chain.

### C. Interplay Between TF and SRP on the Ribosome

The role of TF in protein targeting and folding is not clear. It has been suggested that TF plays a discriminatory role in the targeting process by interacting specifically with the early mature region of secretory proteins, thereby reducing the affinity of their mildly hydrophobic signal peptide for SRP [38, 70]. This would select secretory proteins for subsequent interaction with the downstream chaperone SecB. However, other studies suggest that TF binds to the nascent chain by default and that only SRP can confer targeting specificity because of its high affinity for newly synthesized, strongly hydrophobic targeting sequences in nascent IMPs [83, 84]. Investigations on the molecular contacts of short nascent polypeptides in the *E. coli* cytoplasm emphasizes the fact that interaction of nascent chains with TF near the ribosomal exit site at L23 occurs by default irrespective of the length and nature of the polypeptide chains ([17, 19]; Figure 1.2). The data also indicate that SRP specifically interacts with a hydrophobic TM segment of nascent IMPs, when the first TM segment is barely exposed outside the ribosome, and not with a mildly hydrophobic signal peptide of nascent secretory proteins, effectively contending TF. Apparently, IMPs are selected by SRP for cotranslational targeting at a very early stage during biogenesis [17, 19]. Interestingly, it was found that the N-terminus of nascent IMP chains [19], and artificial TM segment-bearing substrates [18], remain in close contact to the ribosomal exit site during growth of the nascent proteins, suggesting a tethering of the nascent chains to the ribosome, so that they can efficiently interact with factors like the ribosome-bound SRP. When the nascent chains reach a certain length, contact with SRP is lost, giving room for TF to protect the nascent polypeptides against aberrant interactions, until a second TM segment of sufficient hydrophobicity comes along [18, 19]. The second TM segment may provide an

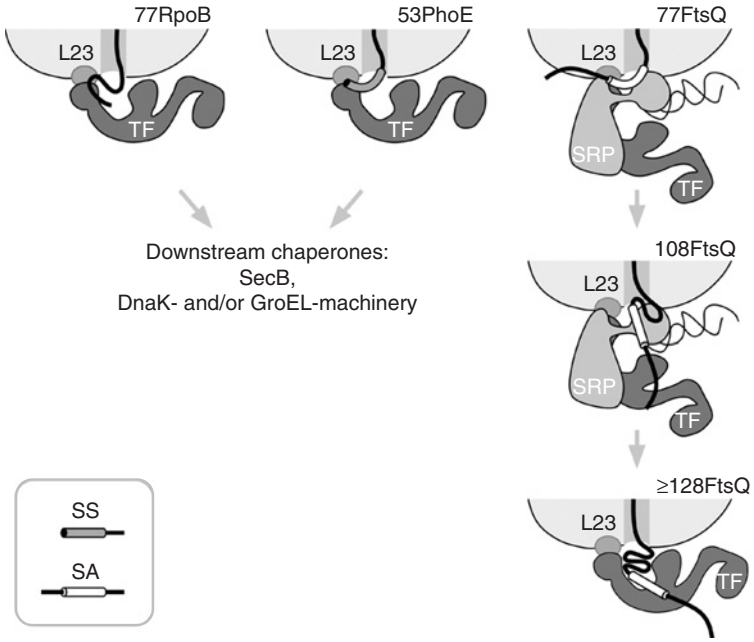



FIG. 1.2. Model for the first contacts of nascent polypeptides in the *E. coli* cytosol. The schemes represent the bottom part of the 50S subunit where the nascent chain leaves the ribosome. The polypeptide exit tunnel is shown in dark gray. A nascent chain is represented by a black solid line and is positioned inside the tunnel. At the exit region of the tunnel the ribosomal protein L23 is depicted in dark gray. TF binds L23 and arches over the exit site. SRP also associates with L23, sometimes simultaneously with TF. SRP is shown here in light gray, comprising Ffh and 4.5S RNA (black colored coiled strand). Most likely, TF must undergo a substantial change in conformation and orientation to allow the SRP to access the nascent IMP. An α-helical signal sequence (SS) is shown as a light gray colored cylinder and an α-helical TM segment/signal anchor (SA) sequence as a white colored cylinder. The nascent chains exit the ribosome in a looped conformation with the N-terminus tethered near L23. Nascent 77RpoB (cytosolic) and 53PhoE (outer membrane protein) exclusively contact L23 and TF. In contrast, 77FtsQ (IMP) also contact SRP. On growth of the nascent IMP, the TM segment is transiently sequestered near the exit site by SRP (right middle scheme; "108FtsQ"). On further elongation, both the TM segment and downstream sequences lose contact with SRP to the benefit of TF (right bottom scheme; "≥128FtsQ"). Most likely, the SRP is released from the ribosome at this stage, giving room to TF. (*Reprinted from* The Journal of Biologycal Chemistry *[19]*.)

additional chance for SRP-mediated targeting as demonstrated for polytopic ER membrane proteins [85, 86], and/or may mediate the contact of Ffh with upstream hydrophobic sequences [18].

The precise interplay between SRP, TF, and emerging nascent chains at the *E. coli* ribosome is not known. TF and SRP can bind simultaneously to ribosomes in *in vitro* experimental systems, which suggests separate binding

sites on L23 [47, 73]. This implies that the orientation of SRP and TF toward the exit site must be flexible and able to respond quickly to passing sequences or sequences still hidden in the ribosome [82]. Structures of homologous ribosome/TF/SRP complexes bearing nascent chains of different nature are needed to gain more insight into the topogenic and conformational changes of all partners near the nascent chain exit site on the ribosome. Note that *in vivo* SRP is at least ~100 times less abundant than TF and ribosomes, and needs to be recruited to ribosomes translating IMPs. It is unclear how TF that is already docked at L23 influences this process.

Cryo-EM studies define the *E. coli* L23 region near the exit site also as the main connection between the ribosome and the SecYEG complex [44]. Both TF and SRP must be released from L23 to allow its transfer to SecY at the membrane. The SRP receptor FtsY might induce dissociation not only of SRP but also of TF. SRP is released in a conserved GTP-dependent reaction on direct intimate contact between Ffh and FtsY (reviewed in [87]). Also TF was released from vacant ribosomes by FtsY in an *in vitro* binding assay [73], although this was not confirmed in a comparable study [47]. Furthermore, it was described that FtsY does not interfere with TF binding to ribosome-associated nascent chains in an *in vitro* cross-linking assay [18]. Hence, it is not yet clear how and when TF is released from L23. Because binding of L23 to TF and SecY seems mutually exclusive, TF could prevent the unproductive association of ribosomes that do not carry a nascent IMP with the Sec translocon by obstructing the L23 component of the ribosome–translocon junction site. Interestingly, inactivation of the TF gene accelerates the translocation of secretory proteins and suppresses their dependence on SecB, whereas overproduction of TF retards their translocation [88]. Whether these effects are related to an altered localization of the translating ribosomes remains to be determined.

### D. THE NASCENT POLYPEPTIDE-ASSOCIATED COMPLEX

It seems relevant to discuss here the nascent polypeptide-associated complex (NAC), a heterodimeric protein complex composed of an α- and a β-subunit, present in archaea and eukaryotes, because it shares properties with bacterial TF [89]. Although NAC has no structural homology to TF, NAC binds like TF to vacant ribosomes and contacts nascent polypeptide chains emerging from translating ribosomes [90, 91]. NAC was shown to bind via the N-terminal segment of the β-subunit to protein L25, the homologue of L23, on the yeast ribosome, where it can interact with nascent polypeptide chains [50, 92]. It was demonstrated that bacterial TF can bind *in vitro* to L25 on the yeast ribosome, where it cross-links to short nascent polypeptide chains translated *in vitro* by yeast ribosomes [93]. Interestingly,

however, TF cannot compensate for the loss of NAC function *in vivo* [50], suggesting divergent functions of NAC and TF. Interestingly, the α-subunit of NAC was shown to interact with SRP [50], which may point to a role for NAC in the protein-targeting process. However, such a role for NAC remains controversial. It has been proposed that NAC influences the fidelity of the cotranslational targeting of nascent chains to the ER membrane [90]. Binding of NAC to signal-less nascent chains would prevent them from interacting with SRP [94]. This was emphasized by the finding that deletion of NAC results in a pronounced growth defect at 37 °C and an induction of the unfolded protein response (UPR) pathway, which is indicative for the accumulation of misfolded proteins in the ER [50]. In addition, binding of NAC to the ribosomal exit of RNCs or ribosomes would prevent their interaction with the protein-conducting channel in the ER membrane [94, 95]. However, other groups found that neither the extent nor the kinetics of ribosome binding to microsomes were altered by the presence or absence of NAC, and they concluded that NAC has no direct role in the targeting process [96, 97]. Clearly, these contradictory views await further investigation to solve the role of NAC, if any, in the protein-targeting process.

### E. The Hsp70–RAC Triad

In addition to NAC, eukaryotic cells possess another major ribosome-associated chaperone system that can interact cotranslationally with nascent chains. This system is composed of two Hsp70 members, namely Ssb and Ssz, and Ssz's Hsp40 cochaperone Zuo (reviewed in [98]). Ssz and Zuo form a complex called ribosome-associated complex (RAC), which stimulates the ATPase activity of Ssb and forms together with Ssb the so-called ribosomal chaperone triad [99]. The ribosomal attachment site for the Ssb/Ssz/Zuo chaperone triad is unknown. Ssb can contact, however, short nascent chains [100], suggesting that the location on the ribosome of the chaperone triad is near the polypeptide exit site. Strikingly, it has been shown that, in spite of the structural dissimilarity with the Ssb/Ssz/Zuo chaperone system, TF can compensate for the loss of the yeast triad *in vivo* [93]. TF can partially restore the cation-sensitive and cold-sensitive phenotype of yeast cells lacking Ssb/Ssz/Zuo [93]. How TF accomplishes this is not known.

## IV. SRP-Mediated Targeting

Signal peptides of target proteins are specifically recognized by SRP as they emerge from the ribosome. On TM segment binding, the RNC–SRP complex is targeted to the Sec translocon through a specific interaction

between SRP and its receptor, FtsY. The transfer of the RNC to SecY at the Sec translocon is coordinated by the combined action of SRP and FtsY (Figure 1.3).

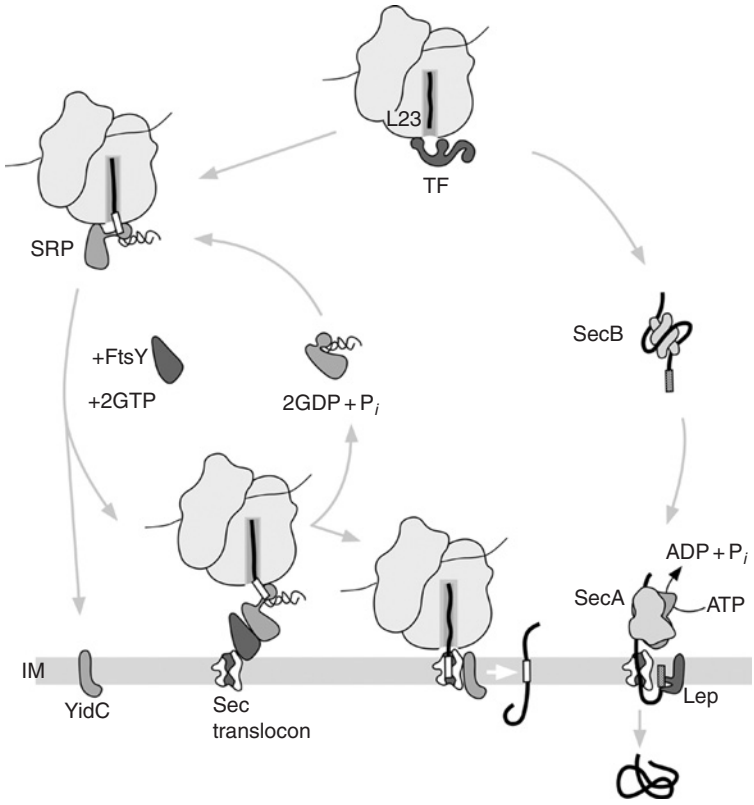



FIG. 1.3. Model of the SRP- and SecB-targeting pathways in *E. coli*. SRP binds to a particularly hydrophobic targeting signal (represented by a white bar) in nascent proteins at L23 near the nascent chain exit site on the large subunit of the ribosome. TF, a cytosolic chaperone that also docks at L23, has a more general affinity for nascent proteins and may influence the interactions of the SRP. FtsY binds to the RNC–SRP complex and supports targeting to the IM through its affinity for SecY and lipids. The nascent protein is transferred to the Sec translocon in a GTP-dependent reaction. The membrane factor YidC mediates IMP assembly both in conjunction with and independently from the Sec translocon. SecB binds to the mature region of presecretory proteins containing a mildly hydrophobic signal peptide (represented by dashed bar), and guides this cargo to the Sec translocon that is identical or similar to the translocon reached via the SRP-pathway. Via a direct interaction with membrane-bound SecA the precursor proteins are relocated to the membrane. During or after translocation, the signal peptide is cleaved off by Leader peptidase (Lep).

## A. General Features of SRP

The SRP was first identified in mammalian cells, in which it targets proteins to the membrane of the ER. The *E. coli* SRP targets proteins to the IM and is one of the simplest known SRPs (Figure 1.4). It consists of one protein (Ffh, SRP54 homologue) [101, 102] and a 4.5S RNA, which is important for Ffh stability [103], and which is probably directly involved in TM segment binding [104]. The most conserved part of the SRP RNA, the domain IV (helix 8 in eukaryotes), binds to Ffh (SRP54). Ffh consists of three domains, the α-helical N-domain, the G-domain (nucleotide-binding domain; GTPase domain), and the C-terminal M-domain (methionine-rich, α-helical domain involved in RNA interaction and signal peptide binding). The N-domain is a four-helix bundle which is tightly connected to the G-domain, forming a distinct structure, commonly referred to as the NG domain [105]. The G-domain shows the classical GTPase fold and contains four conserved sequence motives (GI–GIV). These motifs are involved in nucleotide binding. The SRP GTPases form a distinct subfamily within the superfamily of small GTPases [106].

The mammalian SRP is more complicated and sophisticated in function. It consists of six protein subunits (SRP72, 68, 54, 19, 14, and 9) and is arranged on a 7SL RNA scaffold. SRP can be divided into two functional domains: the Alu domain and the S-domain. SRP14 and SRP9 bind to one end of the 7SL RNA to form the Alu domain, which functions to pause

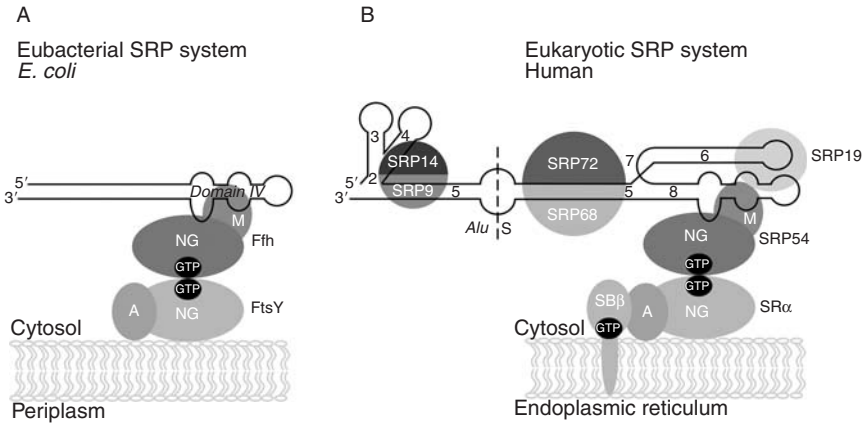



FIG. 1.4. Representation of the *E. coli* and human SRP and SR components (A and B). The proteins are shown in gray and are placed on their binding sites on the SRP RNA. The RNA is shown in black. Structural motifs of the RNA [e.g., helices 2–8 for human SRP RNA (B); domain IV for *E. coli* 4.5S RNA (A)] are marked. (B) A black broken line represents the division of Alu- and S-domain in human SRP. (*Reprinted in adapted form from* Biochimica Et Biophysica Acta *[2]*.)

translation. This feature of SRP is thought to prevent aggregation of fully translated membrane proteins in the cytosol. The remaining four proteins at the other end of the 7SL RNA form the S-domain. It comprises the most conserved part of SRP, namely SRP54 and helix 8 of the 7SL RNA.

It has been shown that the much smaller *E. coli* SRP, which lacks both SRP9 and SRP14, is unable to arrest *in vitro* translation of a heterologous IMP, even though it did bind to its N-terminal TM segment [107]. This is consistent with the notion that SRP-induced translation arrest in eukaryotes is dependent on the Alu domain. However, indirect evidence has been presented suggesting the existence of SRP-mediated translation arrest in *E. coli*, even though the mechanism is still not understood [108].

### B. Signal Peptide Binding to the M-Domain

Signal peptides of target proteins are specifically recognized by SRP as they emerge from the ribosome. Typical signal peptides have a tripartite structure with a 9- to 12-residue-long hydrophobic stretch in the middle [109] that adopts an α-helical conformation. Although for binding to SRP the hydrophobic part of the signal peptide is crucial [83, 84], the flanking regions may also contribute to the interaction ([110]; Section V). On the basis of cross-link and mutagenesis data [111–113], the M-domain of Ffh has been implicated in signal peptide binding. The M-domain is made up of five α-helices and is characterized by an atypically high percentage of methionine residues [113, 114]. The structures of SRP54 from *Sulfolobus solfataricus* alone and in complex with the RNA [115] indeed indicate a binding site for signal peptides, which is in agreement with the binding site suggested from the *E. coli* and *Thermus aquaticus* structures [104, 116]. α-Helices in the M-domain line a hydrophobic groove, which in the absence of a signal peptide is closed to protect the groove from aqueous solvent. On the ribosome, the hydrophobic groove is opened to allow the signal peptide to bind (reviewed in [52]). It has been proposed that the human M-domain forms a slightly different signal peptide-binding site than the one derived from the structures described above (reviewed in [2, 117]), suggesting differences in specificity of signal peptide recognition between eukaryotes and prokaryotes.

### C. The SRP Receptor FtsY

The protein FtsY has been identified as the SRP receptor (SR) in *E. coli* [118]. FtsY is the homologue of SRα in the mammalian SRP receptor, and consist of three domains, the N-terminal acidic domain (A-domain) and the N- and G-domains that are homologues to the N- and G-domains of

protein Ffh. The SRP receptor in mammals consist of two proteins, SRα and SRβ [119–121], which are both GTPases. SRβ serves as a membrane anchor for SRα and is believed to be involved in the signal peptide release from SRP54 [122–124]. In contrast to SRα, *E. coli* FtsY is evenly distributed between the cytosol and the IM [118]. The function of the cytosolic form of FtsY is still unclear, although one study suggests it may be involved in the release of the chaperone TF from the ribosome [73]. Although complex formation of SRP and FtsY already occurs in the cytosol *in vitro*, FtsY requires the context of the IM to ensure faithful transfer of the signal peptide to the translocon [125]. The association of FtsY with the membrane involves both lipids [126, 127] and the translocon component SecY [128]. Both A and NG domains of FtsY have an affinity for the membrane. Intriguingly, evidence from the Bibi laboratory argues that the NG domain expressed alone is fully functional despite less efficient membrane localization [129]. A mechanism by which nascent chain release is coordinated on its interaction with the translocation channel has been suggested. It has been proposed that FtsY binding to the membrane occurs initially through phospholipid binding, followed by targeting to the Sec translocon via an interaction with SecY [127, 128]. The interaction of FtsY with SecY could then induce conformational changes in FtsY to induce the subsequent release of the signal peptide from SRP [128].

### D. Inter- and Intramolecular Communication Between the SRP Components

It has been shown that binding of SRP54 to the ribosome increases the affinity of SRP54 for GTP [130]. Since signal peptide binding and GTP binding occur at distinct domains in SRP54, a communication between the protein domains in the SRP core must exist [131]. A model has been put forward to describe the conformational changes of the SRP core triggered by the interaction with external binding partners, like the RNC and the SRP receptor, and suggest a mechanism for the communication of signal peptide binding in the M-domain to the GTPase domain of SRP54 [52, 131]. Before binding to the ribosome, the ''free'' SRP is likely to adopt a compact structure with the G-domain of SRP54 interacting with RNA helix 8. A direct contact between residues of the M- and N-domains is present, and the M-domain is ''closed'' and incapable of signal peptide binding [115]. In the free state, GTP affinity of SRP54 is low. When bound to a ribosome with or without the presence of a signal peptide, the SRP core adopts the ''open'' conformation observed in the cryo-EM structure [51]. As a signal peptide binds into the hydrophobic groove of SRP54M, a sequence of conformational changes occur in the SRP core. The flexible linker region connecting the M- and G-domain is suggested to play a central role in this

conformational rearrangement. One consequence of the structural changes in all three domains of SRP54 is that SRP is now able to bind GTP with higher affinity that renders SRP competent for targeting of the RNC to the translocon via a specific interaction with the SRP receptor SRα/FtsY (reviewed in [52, 131]).

How are the two GTPases synchronized? GTP binding to SRP54 (Ffh) and SRα (FtsY) is a prerequisite to complex formation, and GTP hydrolysis leads to complex dissociation [132]. The GTPases of Ffh and FtsY stimulate each other on complex formation [133], and therefore, they have been proposed to act as GTPase-activating proteins (GAPs) for each other [134]. Crystal structures of the interacting NG domains of bacterial SRP and SR (Ffh and FtsY from *T. aquaticus*) in the presence of a nonhydrolyzable GTP analogue [135, 136] show that both N-domains rearrange toward the G-domains and that the complex forms an active site at the interface of the two proteins. The individual nucleotide-binding sites are so closely intertwined that the two nucleotides are hydrogen-bonded to each other. On complex formation the NG domains undergo major rearrangements, resulting in conformations suitable for hydrolysis of GTP. However, GTP hydrolysis by SRP and SR must be blocked until the signal peptide is released in order to prevent dissociation of the complex before the RNC is delivered to the translocon. Data suggest that complex formation between SRP and FtsY occurs at the Sec translocon (and not in the cytosol before the RNC has reached the translocon), based on the finding that *in vitro* binding of FtsY to SecY was stabilized by blocking the GTP hydrolysis and was independent on the presence or the absence of SRP [137]. On GTP hydrolysis, the signal peptide is released and the complex dissociates. It is, however, not clear how the release of the signal peptide by the SRP M-domain is communicated to the GTPase domains of both SRP and SR to allow GTP hydrolysis. However, several biochemical studies have shown that the interface between the N- and G-domain may play an important role in the transmission of information between signal peptide binding by the M-domain and the GTPase [138, 139]. In addition, structural studies revealed that there is flexibility between the protein domains in the SRP core, and suggest that the linker region between M- and G-domain is involved in interdomain communication (reviewed in [52, 131]).

## V. Selection of Protein for SRP-Mediated Targeting

In mammalian cells, SRP is the only known dedicated targeting factor to guide both presecretory proteins and membrane proteins to the Sec translocon in the ER membrane [140]. How the bacterial cell selects the

substrates for SRP-mediated targeting is not entirely clear. Considerable evidence indicated that the preferred substrates for targeting by SRP are IMPs (reviewed in [2]). Biogenesis of most IMPs in *E. coli* is strongly hampered at decreased cellular concentrations of SRP components. In addition, genetic screens designed to identify components involved in the biogenesis of IMPs, yielded mutations in all known components of the SRP targeting route [141, 142]. Furthermore, cross-link studies of *in vitro* synthesized short nascent IMPs in homologous systems demonstrated cotranslational recognition of TM segments by SRP [17, 19, 39, 70, 84]. In addition, *in vitro* targeting of nascent IMPs to the IM appeared to be dependent on the presence of SRP [143].

SRP-independent posttranslational targeting or targeting at a late stage during translation in *E. coli* is fulfilled mainly by the SecB pathway (Figure 1.3). The SecB chaperone, which is only present in Gram-negative bacteria, interacts with the mature moiety of secretory proteins containing a cleavable signal peptide, and keeps them in an unfolded, translocation competent state [144]. In addition to maintaining preproteins in a translocation competent conformation, SecB also participates in directing preproteins into the translocation pathway by its specific interaction with the membrane-bound ATPase SecA [145]. After the precursor proteins have been located at the membrane through the SecB–SecA interaction, they are released from SecB and transferred to the SecA–translocon complex for further translocation over the IM [146]. It should be noted that the absence of SecB in *E. coli* does not significantly affect cell growth [147], and many proteins are exported equally well in *secB* mutants (reviewed in [148]). However, the absence of SecB was shown to result in the aggregation of secretory proteins in the cytoplasm, and deletion of *secB* affects the targeting kinetics of secretory proteins [147].

The apparent specificity of the *E. coli* SRP is explained by a preference of Ffh for the relatively hydrophobic targeting signals present in nascent IMPs, as suggested by *in vitro* cross-linking studies [17, 19, 70, 84]. Consistently, a signal peptide of a secretory protein can be recognized by SRP when its hydrophobicity is increased, rerouting the fused passenger proteins via the SRP pathway [83, 149–152]. Interestingly, a single replacement of a helix-breaking glycine residue by a helix-promoting leucine residue in the hydrophobic core of the PhoE signal peptide, a SecB-dependent outer membrane protein, also promoted SRP binding [149]. This suggests that the structure of the signal peptide may be relevant for SRP binding as well. Despite its apparent specificity for IMPs, the *E. coli* SRP may also play a modest role in bacterial protein export. Results obtained from a screen for native *E. coli* signal peptides, which direct fused reporter proteins to the SRP pathway and therefore can enhance cotranslational processing,

suggest that about 10% of all periplasmic proteins are exported via the SRP pathway [153]. One example of a native *E. coli* protein that is recognized by SRP is the periplasmic enzyme DsbA that contains an unusually hydrophobic signal peptide [154].

Hydrophobicity does not appear to be the only key feature for selection by the SRP. Interestingly, it was shown that the bacterial autotransporter Hbp [155] is transported over the IM via the SRP/SecA/SecYEG pathway. Autotransporters contain unusual long but not particularly hydrophobic signal peptides. The C-terminal half of the long signal peptide resembles a classical signal peptide comprising well-defined N, H, and C regions, but with an uncommonly high net positive charge in the N region, which may play a role in SRP RNA-signal peptide binding [110, 155]. The significance of the N-terminal extension of the long signal peptide is not clear but a role in targeting pathway selection has been suggested [156]. The N-terminal extension might rout presecretory proteins into a posttranslational targeting pathway by exerting an inhibitory effect on their interaction with SRP, TF, and the SecYEG translocon [156]. On the other hand, removal of the N-terminus of the autotransporter Hbp had little effect on targeting pathway specificity (Jong and Luirink, unpublished results).

### A. SRP-MEDIATED TARGETING IN *E. COLI* OCCURS COTRANSLATIONALLY

Considerable evidence indicate that SRP-mediated targeting and insertion of IMPs occur in a cotranslational fashion [154, 157]. The affinity of *E. coli* ribosomes for the Sec translocon supports the existence of a cotranslational insertion mechanism in *E. coli* [158]. Moreover, a cryo-EM reconstruction of the dimeric *E. coli* SecYEG channel tethered to a ribosome-nascent FtsQ complex has been reported, again supporting the cotranslational mode of IMP insertion [44]. Also in agreement with the cotranslational targeting of IMPs by SRP, cross-link studies of *in vitro* synthesized nascent IMPs demonstrated recognition of the TM segment in short nascent IMPs by the SRP [17, 19, 39, 70, 84]. When the *in vitro* translation system is supplemented with inverted membrane vesicles, SRP is released from the nascent chain in a reaction that requires FtsY and GTP [125, 159]. On release, the short nascent chain is transferred to the Sec translocon components [70, 125].

Cotranslational targeting of IMPs may start very early during biogenesis when the first TM segments are barely exposed outside the ribosome. It has been shown that interactions with SRP and SecYEG already occur when the first TM segment is not completely exposed outside the ribosome [17]. In principle, this result suggests that nascent IMPs can be handed over from SRP to the translocon at an early stage during biogenesis, as in eukaryotes.

It remains to be established whether *in vivo*, in the absence of translation arrest and given the speed of prokaryotic translation, the translocon is reached at such an early phase in the biogenesis of IMPs.

Independent support for a cotranslational role of SRP was previously presented [150]. A hybrid LamB–LacZ construct that is normally targeted via the SecB pathway and jams the Sec translocon by posttranslational folding of the LacZ moiety was rerouted into the SRP pathway by increasing the hydrophobicity of its signal peptide. The hydrophobic signal peptide provoked SRP-dependent, presumably cotranslational, secretion into the periplasm and thereby prevented lethal jamming of the Sec translocon by the hybrid protein.

Collectively, the data are suggestive of cotranslational targeting and membrane insertion of SRP-dependent IMPs.

### B. How Important Is SRP for Bacteria?

With respect to the different targeting pathways, how crucial is SRP for bacterial cell survival? In *E. coli*, Ffh, 4.5S RNA, and FtsY are essential for growth, which seems in agreement with the essential nature of many of its substrate IMPs (reviewed in [160]). Although both Ffh and 4.5S RNA are not abundant, *E. coli* cells can cope with a drastic reduction even of this limited supply of SRP components, partly because efficient IMP targeting prevents a toxic accumulation of aggregated proteins in the cytoplasm [157]. In reducing levels of SRP, the cells react to this toxic accumulation by increasing the levels of the heat shock regulated proteases Lon and ClpQ [157]. Notably, a substantial fraction of the IMPs inserts correctly under SRP-deficient conditions [161, 162], suggesting alternative targeting mechanisms. The hydrophobic domains in the IMPs might contribute to spontaneous targeting to the Sec translocon. Also, ribosomes may support cotranslational targeting through their affinity for the Sec translocon (section V.A).

Bacteria may differ in their dependence on SRP-mediated targeting. Analysis of the extracellular proteome of *Bacillus subtilis* strains with conditional SRP expression suggests that *Bacillus* SRP is also required for the targeting of most secreted proteins [163, 164]. Notably, signal peptides in *B. subtilis*, a Gram-positive bacterium, are generally longer and more hydrophobic, which is critical for signal peptide recognition by the *B. subtilis* SRP [165]. Consistent with a primary role for the *Bacillus* SRP, a SecB homologue supporting posttranslational targeting is not present in this organism [166].

In *E. coli*, the majority of the membrane proteins seem to be cotranslationally targeted by SRP to the Sec translocon, though this awaits confirmation by proteomic analyses. Subsequently, the nascent IMPs move from the

translocon into the lipid bilayer. In recent years, it has become clear that the IMP YidC not only assists in this process but also mediates the insertion of at least some IMPs independent of the Sec translocon (reviewed in [167–169]. Substrates of this Sec-independent/YidC-dependent pathway include small phage coat proteins and the endogenous subunit c of $F_1F_0$-ATPase (reviewed in [167]). Targeting of the phage M13 procoat and Pf3 coat proteins occurs independent of the SRP [151, 170]. $F_0$c, on the other hand, must be delivered at YidC cotranslationally [171, 172], and may require SRP for proper targeting [172], although this remains controversial [171, 173]. It is unknown whether YidC is at any stage connected with the ribosome during insertion and assembly of $F_0$c. Interestingly, the YidC homologue Oxa1p does bind ribosomes during cotranslational protein insertion in mitochondria [174, 175]. However, YidC was unable to complement for the loss of Oxa1p in mitochondria unless it was fused to the ribosome-binding domain of Oxa1p suggesting that YidC has no affinity for ribosomes [176]. Irrespective of the mechanism, cotranslational targeting to YidC is likely to facilitate the assembly of $F_0$c in the IM and prevent aberrant intra- and intermolecular contacts in the cytoplasm that may lead to aggregation [172].

In *Streptococcus mutans*, the SRP pathway seems to play a less important role in membrane protein targeting. Mutations in *ffh*, *ffs*, or *ftsy*, encoding Ffh, SRP RNA, and FtsY, respectively, resulted in acid sensitivity but not in loss of viability [177, 178]. Surprisingly, elimination of one of the two YidC homologues present in *S. mutans*, YidC2, resulted in a similar stress-sensitive phenotype [178]. Mutants lacking both YidC2 and SRP components were more severely impaired in growth than the single deletions suggesting overlapping and compensatory targeting pathways [178]. This suggests an alternative YidC2-dependent cotranslational targeting mechanism in *S. mutans*, perhaps similar to the role of Oxa1p in cotranslational targeting in mitochondria.

## VI. Concluding Remarks

The genetic and biochemical data concerning the SRP-targeting pathway in *E. coli*, together with the recently available structures of most of the ''players in the field,'' have significantly expanded our knowledge of the functions of these players and of the interplay among these protein complexes. In *E. coli*, SRP binds to the large ribosomal subunit component L23 where it interacts with hydrophobic TM segments in nascent IMPs emerging from the ribosome. On binding of a TM segment by the SRP M-domain, a series of conformational changes in the SRP core are signaled to the SRP G-domain, so that GTP can bind with higher affinity. At the IM,

the RNC–SRP complex associates in a twin-like formation with GTP-loaded FtsY bound to SecYEG and lipids. At the SecYEG translocon, the SRP–FtsY complex dissociates on GTP hydrolysis and the TM segment inserts via SecYEG/YidC into the lipid bilayer. Most IMPs are inserted into the IM via the SRP/SecYEG/YidC pathway. However, recently it has become clear that some IMPs use alternative pathways for IM insertion.

Despite these recent advances, several pressing issues need to be addressed. Among those, additional structural investigations of the interactions between modules of the same species are crucial and need to be carried out. Indeed, small structural differences among protein complexes in the three kingdoms of life may have significant impacts on the conclusions we can draw regarding protein folding and protein targeting. For example, the precise interplay between SRP and the chaperone TF on the *E. coli* ribosome is still largely unclear. More structural data are required to get a better view of the conformation of TF bound to ribosomes carrying a nascent chain and of the complexity of SRP–TF binding at the ribosome in the presence of a nascent IMP. These approaches will undoubtedly shed light on the fascinating relationship between chaperones and targeting factors during the early events of protein targeting to the *E. coli* IM.

## Acknowledgments

R.S.U. is supported by a European Molecular Biology Organization long-term fellowship and a grant from the Swiss National Science Foundation (FN-31-65403) to Costa Georgopoulos.

## References

1. von Heijne, G. (1998). Life and death of a signal peptide. *Nature* 396:111–113.
2. Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: Structure and function revisited. *Biochim Biophys Acta* 1694:17–35.
3. Veenendaal, A.K., van der Does, C., and Driessen, A.J. (2004). The protein-conducting channel SecYEG. *Biochim Biophys Acta* 1694:81–95.
4. Sargent, F., Berks, B.C., and Palmer, T. (2006). Pathfinders and trailblazers: A prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* 254:198–207.
5. Ramakrishnan, V., and Moore, P.B. (2001). Atomic structures at last: the ribosome in 2000. *Curr Opin Struct Biol* 11:144–154.
6. Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* 407:327–339.

7. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* 102:615–623.
8. Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289:905–920.
9. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* 107:679–688.
10. Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289:920–930.
11. Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science* 292:883–896.
12. Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. *Science* 310:827–834.
13. Tenson, T., and Ehrenberg, M. (2002). Regulatory nascent peptides in the ribosomal tunnel. *Cell* 108:591–594.
14. Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* 107:361–372.
15. Gabashvili, I.S., Gregory, S.T., Valle, M., Grassucci, R., Worbs, M., Wahl, M.C., Dahlberg, A.E., and Frank, J. (2001). The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. *Mol Cell* 8:181–188.
16. Jong, W.S., Ten Hagen-Jongman, C.M., Genevaux, P., Brunner, J., Oudega, B., and Luirink, J. (2004). Trigger factor interacts with the signal peptide of nascent Tat substrates but does not play a critical role in Tat-mediated export. *Eur J Biochem* 271:4779–4787.
17. Houben, E.N., Zarivach, R., Oudega, B., and Luirink, J. (2005). Early encounters of a nascent membrane protein: Specificity and timing of contacts inside and outside the ribosome. *J Cell Biol* 170:27–35.
18. Eisner, G., Moser, M., Schafer, U., Beck, K., and Muller, M. (2006). Alternate recruitment of signal recognition particle and trigger factor to the signal sequence of a growing nascent polypeptide. *J Biol Chem* 281:7172–7179.
19. Ullers, R.S., Houben, E.N., Brunner, J., Oudega, B., Harms, N., and Luirink, J. (2006). Sequence-specific interactions of nascent *Escherichia coli* polypeptides with trigger factor and signal recognition particle. *J Biol Chem* 281:13999–14005.
20. Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 406:318–322.
21. Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A., Chapman, M.S., and Frank, J. (2003). Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement. *Cell* 113:789–801.
22. Baram, D., and Yonath, A. (2005). From peptide-bond formation to cotranslational folding: Dynamic, regulatory and evolutionary aspects. *FEBS Lett* 579:948–954.
23. Hardesty, B., and Kramer, G. (2001). Folding of a nascent peptide on the ribosome. *Prog Nucleic Acid Res Mol Biol* 66:41–66.
24. Gilbert, R.J., Fucini, P., Connell, S., Fuller, S.D., Nierhaus, K.H., Robinson, C.V., Dobson, C.M., and Stuart, D.I. (2004). Three-dimensional structures of translating ribosomes by Cryo-EM. *Mol Cell* 14:57–66.

25. Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* 116:725–736.
26. Woolhead, C.A., Johnson, A.E., and Bernstein, H.D. (2006). Translation arrest requires two-way communication between a nascent polypeptide and the ribosome. *Mol Cell* 22:587–598.
27. Berisio, R., Schluenzen, F., Harms, J., Bashan, A., Auerbach, T., Baram, D., and Yonath, A. (2003). Structural insight into the role of the ribosomal tunnel in cellular regulation. *Nat Struct Biol* 10:366–370.
28. Voss, N.R., Gerstein, M., Steitz, T.A., and Moore, P.B. (2006). The geometry of the ribosomal polypeptide exit tunnel. *J Mol Biol* 360:893–906.
29. Nakatogawa, H., and Ito, K. (2004). Intraribosomal regulation of expression and fate of proteins. *Chembiochem* 5:48–51.
30. Mankin, A.S. (2006). Nascent peptide in the ''birth canal'' of the ribosome *Trends Biochem Sci* 31:11–13.
31. Nakatogawa, H., and Ito, K. (2002). The ribosomal exit tunnel functions as a discriminating gate. *Cell* 108:629–636.
32. Mitra, K., Schaffitzel, C., Fabiola, F., Chapman, M.S., Ban, N., and Frank, J. (2006). Elongation arrest by SecM via a cascade of ribosomal RNA rearrangements. *Mol Cell* 22:533–543.
33. Johnson, A.E. (2004). Functional ramifications of FRET-detected nascent chain folding far inside the membrane-bound ribosome. *Biochem Soc Trans* 32:668–672.
34. Liao, S., Lin, J., Do, H., and Johnson, A.E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* 90:31–41.
35. Alder, N.N., and Johnson, A.E. (2004). Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *J Biol Chem* 279:22787–22790.
36. Rapoport, T.A., Goder, V., Heinrich, S.U., and Matlack, K.E. (2004). Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol* 14:568–575.
37. Mitra, K., and Frank, J. (2006). A model for co-translational translocation: Ribosome-regulated nascent polypeptide translocation at the protein-conducting channel. *FEBS Lett* 580:3353–3360.
38. Eisner, G., Koch, H.G., Beck, K., Brunner, J., and Muller, M. (2003). Ligand crowding at a nascent signal sequence. *J Cell Biol* 163:35–44.
39. Ullers, R.S., Houben, E.N., Raine, A., ten Hagen-Jongman, C.M., Ehrenberg, M., Brunner, J., Oudega, B., Harms, N., and Luirink, J. (2003). Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *J Cell Biol* 161:679–684.
40. Gu, S.Q., Peske, F., Wieden, H.J., Rodnina, M.V., and Wintermeyer, W. (2003). The signal recognition particle binds to protein L23 at the peptide exit of the *Escherichia coli* ribosome. *RNA* 9:566–573.
41. Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. *Nature* 419:171–174.
42. Blaha, G., Wilson, D.N., Stoller, G., Fischer, G., Willumeit, R., and Nierhaus, K.H. (2003). Localization of the trigger factor binding site on the ribosomal 50S subunit. *J Mol Biol* 326:887–897.
43. Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., and Akey, C.W. (2000). The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 6:1219–1232.

44. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., III, Ban, N., and Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438:318–324.
45. Walter, P., and Blobel, G. (1983). Subcellular distribution of signal recognition particle and 7SL-RNA determined with polypeptide-specific antibodies and complementary DNA probe. *J Cell Biol* 97:1693–1699.
46. Flanagan, J.J., Chen, J.C., Miao, Y., Shao, Y., Lin, J., Bock, P.E., and Johnson, A.E. (2003). Signal recognition particle binds to ribosome-bound signal sequences with fluorescence-detected subnanomolar affinity that does not diminish as the nascent chain lengthens. *J Biol Chem* 278:18628–18637.
47. Raine, A., Ivanova, N., Wikberg, J.E., and Ehrenberg, M. (2004). Simultaneous binding of trigger factor and signal recognition particle to the *E. coli* ribosome. *Biochimie* 86:495–500.
48. Ogg, S.C., and Walter, P. (1995). SRP samples nascent chains for the presence of signal sequences by interacting with ribosomes at a discrete step during translation elongation. *Cell* 81:1075–1084.
49. Pool, M.R., Stumm, J., Fulga, T.A., Sinning, I., and Dobberstein, B. (2002). Distinct modes of signal recognition particle interaction with the ribosome. *Science* 297:1345–1348.
50. Grallath, S., Schwarz, J.P., Bottcher, U.M., Bracher, A., Hartl, F.U., and Siegers, K. (2006). L25 functions as a conserved ribosomal docking site shared by nascent chain-associated complex and signal-recognition particle. *EMBO Rep* 7:78–84.
51. Halic, M., Becker, T., Pool, M.R., Spahn, C.M., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427:808–814.
52. Wild, K., Halic, M., Sinning, I., and Beckmann, R. (2004a). SRP meets the ribosome. *Nat Struct Mol Biol* 11:1049–1053.
53. Lill, R., Crooke, E., Guthrie, B., and Wickner, W. (1988). The ''trigger factor cycle'' includes ribosomes, presecretory proteins, and the plasma membrane *Cell* 54:1013–1018.
54. Crooke, E., and Wickner, W. (1987). Trigger factor: A soluble protein that folds pro-OmpA into a membrane-assembly-competent form. *Proc Natl Acad Sci USA* 84:5216–5220.
55. Stoller, G., Rucknagel, K.P., Nierhaus, K.H., Schmid, F.X., Fischer, G., and Rahfeld, J.U. (1995). A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J* 14:4939–4948.
56. Huang, G.C., Li, Z.Y., Zhou, J.M., and Fischer, G. (2000). Assisted folding of D-glyceraldehyde-3-phosphate dehydrogenase by trigger factor. *Protein Sci* 9:1254–1261.
57. Scholz, C., Stoller, G., Zarnt, T., Fischer, G., and Schmid, F.X. (1997). Cooperation of enzymatic and chaperone functions of trigger factor in the catalysis of protein folding. *EMBO J* 16:54–58.
58. Liu, C.P., and Zhou, J.M. (2004). Trigger factor-assisted folding of bovine carbonic anhydrase II. *Biochem Biophys Res Commun* 313:509–515.
59. Kramer, G., Patzelt, H., Rauch, T., Kurz, T.A., Vorderwulbecke, S., Bukau, B., and Deuerling, E. (2004). Trigger factor peptidyl-prolyl cis/trans isomerase activity is not essential for the folding of cytosolic proteins in *Escherichia coli*. *J Biol Chem* 279:14165–14170.
60. Patzelt, H., Rudiger, S., Brehmer, D., Kramer, G., Vorderwulbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., Bukau, B., and Deuerling, E. (2001). Binding specificity of *Escherichia coli* trigger factor. *Proc Natl Acad Sci USA* 98:14244–14249.
61. Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F.U. (1999). Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* 97:755–765.

62. Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* 400:693–696.
63. Deuerling, E., Patzelt, H., Vorderwulbecke, S., Rauch, T., Kramer, G., Schaffitzel, E., Mogk, A., Schulze-Specking, A., Langen, H., and Bukau, B. (2003). Trigger Factor and DnaK possess overlapping substrate pools and binding specificities. *Mol Microbiol* 47:1317–1328.
64. Agashe, V.R., Guha, S., Chang, H.C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: Increase in yield at the expense of folding speed. *Cell* 117:199–209.
65. Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep* 5:195–200.
66. Vorderwulbecke, S., Kramer, G., Merz, F., Kurz, T.A., Rauch, T., Zachmann-Brand, B., Bukau, B., and Deuerling, E. (2004). Low temperature or GroEL/ES overproduction permits growth of *Escherichia coli* cells lacking trigger factor and DnaK. *FEBS Lett* 559:181–187.
67. Scholz, C., Mucke, M., Rape, M., Pecht, A., Pahl, A., Bang, H., and Schmid, F.X. (1998). Recognition of protein substrates by the prolyl isomerase trigger factor is independent of proline residues. *J Mol Biol* 277:723–732.
68. Hesterkamp, T., Hauser, S., Lutcke, H., and Bukau, B. (1996). *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. *Proc Natl Acad Sci USA* 93:4437–4441.
69. Valent, Q.A., Kendall, D.A., High, S., Kusters, R., Oudega, B., and Luirink, J. (1995). Early events in preprotein recognition in *E. coli*: Interaction of SRP and trigger factor with nascent polypeptides. *EMBO J* 14:5494–5505.
70. Beck, K., Wu, L.F., Brunner, J., and Muller, M. (2000). Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J* 19:134–143.
71. Maier, R., Scholz, C., and Schmid, F.X. (2001). Dynamic association of trigger factor with protein substrates. *J Mol Biol* 314:1181–1190.
72. Maier, R., Eckert, B., Scholz, C., Lilie, H., and Schmid, F.X. (2003). Interaction of trigger factor with the ribosome. *J Mol Biol* 326:585–592.
73. Buskiewicz, I., Deuerling, E., Gu, S.Q., Jockel, J., Rodnina, M.V., Bukau, B., and Wintermeyer, W. (2004). Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. *Proc Natl Acad Sci USA* 101:7902–7906.
74. Raine, A., Lovmar, M., Wikberg, J., and Ehrenberg, M. (2006). Trigger factor binding to ribosomes with nascent peptide chains of varying lengths and sequences. *J Biol Chem* 281:28033–28038.
75. Hesterkamp, T., and Bukau, B. (1996). The *Escherichia coli* trigger factor. *FEBS Lett* 389:32–34.
76. Stoller, G., Tradler, T., Rucknagel, K.P., Rahfeld, J.U., and Fischer, G. (1996). An 11.8 kDa proteolytic fragment of the *E. coli* trigger factor represents the domain carrying the peptidyl-prolyl cis/trans isomerase activity. *FEBS Lett* 384:117–122.
77. Hesterkamp, T., Deuerling, E., and Bukau, B. (1997). The amino-terminal 118 amino acids of *Escherichia coli* trigger factor constitute a domain that is necessary and sufficient for binding to ribosomes. *J Biol Chem* 272:21865–21871.

78. Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431:590–596.
79. Hoffmann, A., Merz, F., Rutkowska, A., Zachmann-Brand, B., Deuerling, E., and Bukau, B. (2006). Trigger factor forms a protective shield for nascent polypeptides at the ribosome. *J Biol Chem* 281:6539–6545.
80. Tomic, S., Johnson, A.E., Hartl, F.U., and Etchells, S.A. (2006). Exploring the capacity of trigger factor to function as a shield for ribosome bound polypeptide chains. *FEBS Lett* 580:72–76.
81. Yonath, A. (2005). Ribosomal crystallography: Peptide bond formation, chaperone assistance and antibiotics activity. *Mol Cells* 20:1–16.
82. Schlunzen, F., Wilson, D.N., Tian, P., Harms, J.M., McInnes, S.J., Hansen, H.A., Albrecht, R., Buerger, J., Wilbanks, S.M., and Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: Implications for protein folding and SRP interaction. *Structure (Camb)* 13:1685–1694.
83. Lee, H.C., and Bernstein, H.D. (2001). The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc Natl Acad Sci USA* 98:3471–3476.
84. Valent, Q.A., de Gier, J.W., von Heijne, G., Kendall, D.A., ten Hagen-Jongman, C.M., Oudega, B., and Luirink, J. (1997). Nascent membrane and presecretory proteins synthesized in *Escherichia coli* associate with signal recognition particle and trigger factor. *Mol Microbiol* 25:53–64.
85. Kuroiwa, T., Sakaguchi, M., Omura, T., and Mihara, K. (1996). Reinitiation of protein translocation across the endoplasmic reticulum membrane for the topogenesis of multispanning membrane proteins. *J Biol Chem* 271:6423–6428.
86. Lu, Y., Xiong, X., Helm, A., Kimani, K., Bragin, A., and Skach, W.R. (1998). Co- and posttranslational translocation mechanisms direct cystic fibrosis transmembrane conductance regulator N terminus transmembrane assembly. *J Biol Chem* 273:568–576.
87. Shan, S.O., and Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. *FEBS Lett* 579:921–926.
88. Lee, H.C., and Bernstein, H.D. (2002). Trigger factor retards protein export in *E. coli*. *J Biol Chem* 277:43527–43535.
89. Bukau, B., Deuerling, E., Pfund, C., and Craig, E.A. (2000). Getting newly synthesized proteins into shape. *Cell* 101:119–122.
90. Wiedmann, B., Sakai, H., Davis, T.A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. *Nature* 370:434–440.
91. Shi, X., Parthun, M.R., and Jaehning, J.A. (1995). The yeast EGD2 gene encodes a homologue of the alpha NAC subunit of the human nascent-polypeptide-associated complex. *Gene* 165:199–202.
92. Wegrzyn, R.D., Hofmann, D., Merz, F., Nikolay, R., Rauch, T., Graf, C., and Deuerling, E. (2006). A conserved motif is prerequisite for the interaction of NAC with ribosomal protein L23 and nascent chains. *J Biol Chem* 281:2847–2857.
93. Rauch, T., Hundley, H.A., Pfund, C., Wegrzyn, R.D., Walter, W., Kramer, G., Kim, S.Y., Craig, E.A., and Deuerling, E. (2005). Dissecting functional similarities of ribosome-associated chaperones from Saccharomyces cerevisiae and *Escherichia coli*. *Mol Microbiol* 57:357–365.
94. Lauring, B., Kreibich, G., and Weidmann, M. (1995a). The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. *Proc Natl Acad Sci USA* 92:9435–9439.

95. Lauring, B., Sakai, H., Kreibich, G., and Wiedmann, M. (1995b). Nascent polypeptide-associated complex protein prevents mistargeting of nascent chains to the endoplasmic reticulum. *Proc Natl Acad Sci USA* 92:5411–5415.
96. Neuhof, A., Rolls, M.M., Jungnickel, B., Kalies, K.U., and Rapoport, T.A. (1998). Binding of signal recognition particle gives ribosome/nascent chain complexes a competitive advantage in endoplasmic reticulum membrane interaction. *Mol Biol Cell* 9:103–115.
97. Raden, D., and Gilmore, R. (1998). Signal recognition particle-dependent targeting of ribosomes to the rough endoplasmic reticulum in the absence and presence of the nascent polypeptide-associated complex. *Mol Biol Cell* 9:117–130.
98. Wegrzyn, R.D., and Deuerling, E. (2005). Molecular guardians for newborn proteins: Ribosome-associated chaperones and their role in protein folding. *Cell Mol Life Sci* 62:2727–2738.
99. Gautschi, M., Lilie, H., Funfschilling, U., Mun, A., Ross, S., Lithgow, T., Rucknagel, P., and Rospert, S. (2001). RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. *Proc Natl Acad Sci USA* 98:3762–3767.
100. Pfund, C., Lopez-Hoyo, N., Ziegelhoffer, T., Schilke, B.A., Lopez-Buesa, P., Walter, W.A., Wiedmann, M., and Craig, E.A. (1998). The molecular chaperone Ssb from Saccharomyces cerevisiae is a component of the ribosome-nascent chain complex. *EMBO J* 17:3981–3989.
101. Romisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M., and Dobberstein, B. (1989). Homology of 54K protein of signal-recognition particle, docking protein and two *E. coli* proteins with putative GTP-binding domains. *Nature* 340:478–482.
102. Bernstein, H.D., Poritz, M.A., Strub, K., Hoben, P.J., Brenner, S., and Walter, P. (1989). Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature* 340:482–486.
103. Jensen, C.G., and Pedersen, S. (1994). Concentrations of 4.5S RNA and Ffh protein in *Escherichia coli*: The stability of Ffh protein is dependent on the concentration of 4.5S RNA. *J Bacteriol* 176:7148–7154.
104. Batey, R.T., Rambo, R.P., Lucast, L., Rha, B., and Doudna, J.A. (2000). Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* 287:1232–1239.
105. Freymann, D.M., Keenan, R.J., Stroud, R.M., and Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature* 385:361–364.
106. Bourne, H.R., Sanders, D.A., and McCormick, F. (1990). The GTPase superfamily: A conserved switch for diverse cell functions. *Nature* 348:125–132.
107. Raine, A., Ullers, R., Pavlov, M., Luirink, J., Wikberg, J.E., and Ehrenberg, M. (2003). Targeting and insertion of heterologous membrane proteins in *E. coli*. *Biochimie* 85:659–668.
108. Avdeeva, O.N., Myasnikov, A.G., Sergiev, P.V., Bogdanov, A.A., Brimacombe, R., and Dontsova, O.A. (2002). Construction of the 'minimal' SRP that interacts with the translating ribosome but not with specific membrane receptors in *Escherichia coli*. *FEBS Lett* 514:70–73.
109. von Heijne, G. (1985). Signal sequences. The limits of variation. *J Mol Biol* 184:99–105.
110. Peterson, J.H., Woolhead, C.A., and Bernstein, H.D. (2003). Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. *J Biol Chem* 278:46155–46162.
111. Lutcke, H., High, S., Romisch, K., Ashford, A.J., and Dobberstein, B. (1992). The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO J* 11:1543–1551.
112. Romisch, K., Ribes, V., High, S., Lutcke, H., Tollervey, D., and Dobberstein, B. (1990a). Structure and function of signal recognition particle (SRP). *Mol Biol Rep* 14:71–72.

113. Zopf, D., Bernstein, H.D., Johnson, A.E., and Walter, P. (1990). The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J* 9:4511–4517.
114. Romisch, K., Webb, J., Lingelbach, K., Gausepohl, H., and Dobberstein, B. (1990b). The 54-kD protein of signal recognition particle contains a methionine-rich RNA binding domain. *J Cell Biol* 111:1793–1802.
115. Rosendal, K.R., Wild, K., Montoya, G., and Sinning, I. (2003). Crystal structure of the complete core of archaeal signal recognition particle and implications for interdomain communication. *Proc Natl Acad Sci USA* 100:14701–14706.
116. Keenan, R.J., Freymann, D.M., Walter, P., and Stroud, R.M. (1998). Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell* 94:181–191.
117. Doudna, J.A., and Batey, R.T. (2004). Structural insights into the signal recognition particle. *Annu Rev Biochem* 73:539–557.
118. Luirink, J., ten Hagen-Jongman, C.M., van der Weijden, C.C., Oudega, B., High, S., Dobberstein, B., and Kusters, R. (1994). An alternative protein targeting pathway in *Escherichia coli*: Studies on the role of FtsY. *EMBO J* 13:2289–2296.
119. Meyer, D.I., and Dobberstein, B. (1980). A membrane component essential for vectorial translocation of nascent proteins across the endoplasmic reticulum: requirements for its extraction and reassociation with the membrane. *J Cell Biol* 87:498–502.
120. Gilmore, R., Walter, P., and Blobel, G. (1982). Protein translocation across the endoplasmic reticulum.II. Isolation and characterization of the signal recognition particle receptor. *J Cell Biol* 95:470–477.
121. Tajima, S., Lauffer, L., Rath, V.L., and Walter, P. (1986). The signal recognition particle receptor is a complex that contains two distinct polypeptide chains. *J Cell Biol* 103:1167–1178.
122. Song, W., Raden, D., Mandon, E.C., and Gilmore, R. (2000). Role of Sec61alpha in the regulated transfer of the ribosome-nascent chain complex from the signal recognition particle to the translocation channel. *Cell* 100:333–343.
123. Fulga, T.A., Sinning, I., Dobberstein, B., and Pool, M.R. (2001). SRbeta coordinates signal sequence release from SRP with ribosome binding to the translocon. *EMBO J* 20:2338–2347.
124. Schwartz, T., and Blobel, G. (2003). Structural basis for the function of the beta subunit of the eukaryotic signal recognition particle receptor. *Cell* 112:793–803.
125. Valent, Q.A., Scotti, P.A., High, S., de Gier, J.W., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998). The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J* 17:2504–2512.
126. de Leeuw, E., te Kaat, K., Moser, C., Menestrina, G., Demel, R., de Kruijff, B., Oudega, B., Luirink, J., and Sinning, I. (2000). Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity. *EMBO J* 19:531–541.
127. Millman, J.S., Qi, H.Y., Vulcu, F., Bernstein, H.D., and Andrews, D.W. (2001). FtsY binds to the *Escherichia coli* inner membrane via interactions with phosphatidylethanolamine and membrane proteins. *J Biol Chem* 276:25982–25989.
128. Angelini, S., Deitermann, S., and Koch, H.G. (2005). FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. *EMBO Rep* 6:476–481.
129. Eitan, A., and Bibi, E. (2004). The core *Escherichia coli* signal recognition particle receptor contains only the N and G domains of FtsY. *J Bacteriol* 186:2492–2494.
130. Bacher, G., Lutcke, H., Jungnickel, B., Rapoport, T.A., and Dobberstein, B. (1996). Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature* 381:248–251.

131. Wild, K., Rosendal, K.R., and Sinning, I. (2004b). A structural step into the SRP cycle. *Mol Microbiol* 53:357–363.
132. Connolly, T., Rapiejko, P.J., and Gilmore, R. (1991). Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. *Science* 252:1171–1173.
133. Miller, J.D., Bernstein, H.D., and Walter, P. (1994). Interaction of *E. coli* Ffh/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and its receptor. *Nature* 367:657–659.
134. Powers, T., and Walter, P. (1995). Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science* 269:1422–1424.
135. Focia, P.J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004). Heterodimeric GTPase core of the SRP targeting complex. *Science* 303:373–377.
136. Egea, P.F., Shan, S.O., Napetschnig, J., Savage, D.F., Walter, P., and Stroud, R.M. (2004). Substrate twinning activates the signal recognition particle and its receptor. *Nature* 427:215–221.
137. Angelini, S., Boy, D., Schiltz, E., and Koch, H.G. (2006). Membrane binding of the bacterial signal recognition particle receptor involves two distinct binding sites. *J Cell Biol* 174:715–724.
138. Lu, Y., Qi, H.Y., Hyndman, J.B., Ulbrandt, N.D., Teplyakov, A., Tomasevic, N., and Bernstein, H.D. (2001). Evidence for a novel GTPase priming step in the SRP protein targeting pathway. *EMBO J* 20:6724–6734.
139. Newitt, J.A., and Bernstein, H.D. (1997). The N-domain of the signal recognition particle 54-kDa subunit promotes efficient signal sequence binding. *Eur J Biochem* 245:720–729.
140. Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. *Annu Rev Biochem* 70:755–775.
141. Tian, H., Boyd, D., and Beckwith, J. (2000). A mutant hunt for defects in membrane protein assembly yields mutations affecting the bacterial signal recognition particle and Sec machinery. *Proc Natl Acad Sci USA* 97:4730–4735.
142. Tian, H., and Beckwith, J. (2002). Genetic screen yields mutations in genes encoding all known components of the *Escherichia coli* signal recognition particle pathway. *J Bacteriol* 184:111–118.
143. Koch, H.G., Hengelage, T., Neumann-Haefelin, C., MacFarlane, J., Hoffschulte, H.K., Schimz, K.L., Mechler, B., and Muller, M. (1999). *In vitro* studies with purified components reveal signal recognition particle (SRP) and SecA/SecB as constituents of two independent protein-targeting pathways of *Escherichia coli*. *Mol Biol Cell* 10:2163–2173.
144. Lecker, S.H., Driessen, A.J., and Wickner, W. (1990). ProOmpA contains secondary and tertiary structure prior to translocation and is shielded from aggregation by association with SecB protein. *EMBO J* 9:2309–2314.
145. Fekkes, P., de Wit, J.G., van der Wolk, J.P., Kimsey, H.H., Kumamoto, C.A., and Driessen, A.J. (1998). Preprotein transfer to the *Escherichia coli* translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol Microbiol* 29:1179–1190.
146. Fekkes, P., van der Does, C., and Driessen, A.J. (1997). The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J* 16:6105–6113.
147. Baars, L., Ytterberg, A.J., Drew, D., Wagner, S., Thilo, C., van Wijk, K.J., and de Gier, J.W. (2006). Defining the role of the *Escherichia coli* chaperone SecB using comparative proteomics. *J Biol Chem* 281:10024–10034.
148. Randall, L.L., and Hardy, S.J. (2002). SecB, one small chaperone in the complex milieu of the cell. *Cell Mol Life Sci* 59:1617–1623.

149. Adams, H., Scotti, P.A., De Cock, H., Luirink, J., and Tommassen, J. (2002). The presence of a helix breaker in the hydrophobic core of signal sequences of secretory proteins prevents recognition by the signal-recognition particle in *Escherichia coli*. *Eur J Biochem* 269:5564–5571.
150. Bowers, C.W., Lau, F., and Silhavy, T.J. (2003). Secretion of LamB-LacZ by the signal recognition particle pathway of *Escherichia coli*. *J Bacteriol* 185:5697–5705.
151. de Gier, J.W., Scotti, P.A., Saaf, A., Valent, Q.A., Kuhn, A., Luirink, J., and von Heijne, G. (1998). Differential use of the signal recognition particle translocase targeting pathway for inner membrane protein assembly in *Escherichia coli*. *Proc Natl Acad Sci USA* 95:14646–14651.
152. Kim, J., Rusch, S., Luirink, J., and Kendall, D.A. (2001). Is Ffh required for export of secretory proteins? *FEBS Lett* 505:245–248.
153. Huber, D., Boyd, D., Xia, Y., Olma, M.H., Gerstein, M., and Beckwith, J. (2005). Use of thioredoxin as a reporter to identify a subset of *Escherichia coli* signal sequences that promote signal recognition particle-dependent translocation. *J Bacteriol* 187:2983–2991.
154. Schierle, C.F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. (2003). The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway. *J Bacteriol* 185:5706–5713.
155. Sijbrandi, R., Urbanus, M.L., ten Hagen-Jongman, C.M., Bernstein, H.D., Oudega, B., Otto, B.R., and Luirink, J. (2003). Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J Biol Chem* 278:4654–4659.
156. Peterson, J.H., Szabady, R.L., and Bernstein, H.D. (2006). An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor, and the SecYEG complex. *J Biol Chem* 281:9038–9048.
157. Bernstein, H.D., and Hyndman, J.B. (2001). Physiological basis for conservation of the signal recognition particle targeting pathway in *Escherichia coli*. *J Bacteriol* 183:2187–2197.
158. Prinz, A., Behrens, C., Rapoport, T.A., Hartmann, E., and Kalies, K.U. (2000). Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. *EMBO J* 19:1900–1906.
159. Neumann-Haefelin, C., Schafer, U., Muller, M., and Koch, H.G. (2000). SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. *EMBO J* 19:6419–6426.
160. de Gier, J.W., and Luirink, J. (2001). Biogenesis of inner membrane proteins in *Escherichia coli*. *Mol Microbiol* 40:314–322.
161. Ulbrandt, N.D., Newitt, J.A., and Bernstein, H.D. (1997). The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* 88:187–196.
162. Newitt, J.A., Ulbrandt, N.D., and Bernstein, H.D. (1999). The structure of multiple polypeptide domains determines the signal recognition particle targeting requirement of *Escherichia coli* inner membrane proteins. *J Bacteriol* 181:4561–4567.
163. Hirose, I., Sano, K., Shioda, I., Kumano, M., Nakamura, K., and Yamane, K. (2000). Proteome analysis of Bacillus subtilis extracellular proteins: a two-dimensional protein electrophoretic study. *Microbiology* 146(Pt. 1):65–75.
164. Zanen, G., Antelmann, H., Meima, R., Jongbloed, J.D., Kolkman, M., Hecker, M., van Dijl, J.M., and Quax, W.J. (2006). Proteomic dissection of potential signal recognition particle dependence in protein secretion by *Bacillus subtilis*. *Proteomics* 6:3636–3648.

165. Zanen, G., Houben, E.N., Meima, R., Tjalsma, H., Jongbloed, J.D., Westers, H., Oudega, B., Luirink, J., van Dijl, J.M., and Quax, W.J. (2005). Signal peptide hydrophobicity is critical for early stages in protein export by Bacillus subtilis. *FEBS J* 272:4617–4630.
166. Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S., and van Dijl, J.M. (2000). Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* 64:515–547.
167. Dalbey, R.E., and Kuhn, A. (2004). YidC family members are involved in the membrane insertion, lateral integration, folding, and assembly of membrane proteins. *J Cell Biol* 166:769–774.
168. van der Laan, M., Nouwen, N.P., and Driessen, A.J. (2005). YidC—an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr Opin Microbiol* 8:182–187.
169. Luirink, J., von Heijne, G., Houben, E., and de Gier, J.W. (2005). Biogenesis of inner membrane proteins in *Escherichia coli*. *Annu Rev Microbiol* 59:329–355.
170. Chen, M., Samuelson, J.C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R.E. (2002). Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion. *J Biol Chem* 277:7670–7675.
171. van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A.J. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J Cell Biol* 165:213–222.
172. van Bloois, E., Jan Haan, G., de Gier, J.W., Oudega, B., and Luirink, J. (2004). F(1)F(0) ATP synthase subunit c is targeted by the SRP to YidC in the *E. coli* inner membrane. *FEBS Lett* 576:97–100.
173. Yi, L., Celebi, N., Chen, M., and Dalbey, R.E. (2004). Sec/SRP requirements and energetics of membrane insertion of subunits a, b, and c of the *Escherichia coli* F1F0 ATP synthase. *J Biol Chem* 279:39260–39267.
174. Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and Stuart, R.A. (2003). Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. *EMBO J* 22:6438–6447.
175. Ott, M., Prestele, M., Bauerschmitt, H., Funes, S., Bonnefoy, N., and Herrmann, J.M. (2006). Mba1, a membrane-associated ribosome receptor in mitochondria. *EMBO J* 25:1603–1610.
176. Preuss, M., Ott, M., Funes, S., Luirink, J., and Herrmann, J.M. (2005). Evolution of mitochondrial oxa proteins from bacterial YidC. Inherited and acquired functions of a conserved protein insertion machinery. *J Biol Chem* 280:13004–13011.
177. Crowley, P.J., Svensater, G., Snoep, J.L., Bleiweis, A.S., and Brady, L.J. (2004). An ffh mutant of Streptococcus mutans is viable and able to physiologically adapt to low pH in continuous culture. *FEMS Microbiol Lett* 234:315–324.
178. Hasona, A., Crowley, P.J., Levesque, C.M., Mair, R.W., Cvitkovitch, D.G., Bleiweis, A.S., and Brady, L.J. (2005). Streptococcal viability and diminished stress tolerance in mutants lacking the signal recognition particle pathway or YidC2. *Proc Natl Acad Sci USA* 102:17466–17471.

# 2

# *Sec Protein-Conducting Channel and SecA*

ELI O. VAN DER SLUIS • NICO NOUWEN • ARNOLD J.M. DRIESSEN

*Department of Molecular Microbiology*
*Groningen Biomolecular Sciences and Biotechnology Institute*
*University of Groningen*
*9751 NN Haren, The Netherlands*

## I. Abstract

The Sec machinery facilitates protein translocation membrane insertion and into biological membranes of organisms from all three domains of life. The mechanism of the cotranslational mode of translocation is conserved across the domains, whereas the components involved in posttranslational translocation differ. In addition, significant differences are observed in the composition of the Sec machinery within the bacterial domain. Here, we will review these differences in an evolutionary context, and discuss the latest insights into the structure and dynamics of the translocon and the bacterial motor protein SecA, with emphasis on their oligomeric state(s) during protein translocation.

## II. Introduction

Every cell contains at least one membrane that separates the cytoplasm from the extracellular environment and its intracellular organelles. Embedded within these membranes is a variety of different transport systems that selectively allow passage of molecules, thereby enabling the cell to carefully



DOI: 10.1016/S1874-6047(07)25002-4

control the (bio)chemical composition on both sides of the membrane. Proteins are the largest and most complex molecules that are transported across membranes, and several different transport systems exist that can handle this class of substrates. The Sec machinery is the only protein transport system that is conserved across all three domains of life. It enables protein translocation across the cytoplasmic membrane of bacteria and archaea, the endoplasmic reticulum (ER) membrane of eukarya, and the thylakoid membrane of photosynthetic eukarya [1].

A property that distinguishes the Sec machinery from other transport systems is its ability to transport substrates toward two different cellular compartments: the aqueous environment on the trans side of the membrane or the hydrophobic environment of the membrane itself. In line with that property, the spectrum of substrates that is transported by the Sec machinery ranges from highly hydrophobic to highly hydrophilic proteins. The only feature that all substrates have in common is a hydrophobic N-terminal signal sequence or a transmembrane segment (membrane anchor signal) that ensures substrate recognition and initiation of the translocation process. Most signal sequences are cleaved off by a signal peptidase to convert the preprotein into the mature form, whereas N-terminal transmembrane segments remain attached to the substrate.

The most conserved part of the Sec machinery is the ''translocon,'' a membrane integrated channel that allows the passage of the (pre)proteins across the hydrophobic lipid bilayer [2]. All translocons consist of three evolutionarily related subunits, but nevertheless archaeal and eukaryotic translocons can be distinguished from bacterial and thylakoid translocons on the basis of their amino acid sequences [3]. The translocon can associate with different partners to mediate two conceptually different modes of protein translocation: cotranslational and posttranslational translocation. The first is mainly employed for the insertion of integral membrane proteins (IMPs), and the latter mainly for translocation of secretory proteins [4]. Cotranslational translocation requires the translocon to associate with the ribosome, allowing a direct coupling between synthesis and translocation of the (pre)protein [5]. This process is conserved in all domains of life [6] and driven by ongoing protein synthesis at the ribosome. To prevent synthesis of membrane proteins in the cytoplasm, ribosome nascent chain complexes (RNCs) are targeted to the translocon via the signal recognition particle (SRP) in conjunction with its membrane-bound receptor (SR) [7]. In eukaryotes, protein synthesis is slowed down or arrested until the nascent chain has been transferred from SRP to the translocon [8]. For more details on the mechanism of SRP-dependent targeting, the reader is referred to one of the reviews that have appeared [9, 10].

Posttranslational protein translocation occurs by definition after protein synthesis has been completed and requires the translocon to associate with

a motor protein to provide the driving force for the translocation reaction. In this mode of translocation, the Sec machineries in the various domains of life differ substantially from each other. Posttranslational translocation in bacteria and chloroplasts is driven by the *cis*-acting ATPase SecA [11], whereas in ER membranes it is driven by a *trans*-acting Hsp70-like ATPase termed BiP or Kar2 [12]. Given this topological difference, the molecular mechanism underlying posttranslational translocation is expected to differ largely between the ER and the bacterial cytoplasmic membrane. Post-translational protein translocation has also been suggested to occur in archaea, but these organisms lack a SecA homologue and no apparent energy source is available for a trans-acting motor protein.

## III. Outline

The overall mechanisms of the two modes of protein translocation have been unraveled by groundbreaking studies in the early nineties, employing reconstituted systems from *Escherichia coli* and *Saccharomyces cerevisiae*. The last 5 years have led to a tremendous increase in our insights into the structural basis of protein translocation through the elucidation of high-resolution crystal structures from individual components [13–17] and low- to medium-resolution electron microscopy (EM) structures of a variety of functional complexes [18–21]. These structural and biochemical data have yielded detailed insights into the molecular mechanism underlying protein translocation. This review presents an overview of our current understanding of the structural dynamics of the bacterial Sec machinery during protein translocation. We will focus on conformational changes that occur within the translocon, how they might be induced by (pre) proteins, the ribosome or SecA, and we will highlight major unresolved questions. Some of these issues have received considerable attention in reviews [2, 22–24], and therefore additional emphasis will be on two issues that have not been addressed extensively, that is variations that are observed between Sec machineries of different bacteria and the controversy concerning the oligomeric state(s) of the translocon and SecA during protein translocation.

## IV. Variation and Evolution of the Sec Machinery

### A. The Canonical Bacterial Sec Machinery

In addition to the motor protein SecA and the three translocon proteins (SecY, SecE, and SecG), the Sec machinery of the vast majority of bacteria consists of YidC, SecD, SecF, and YajC. YidC is involved in the insertion of

IMPs into the lipid bilayer by contacting the transmembrane segments of nascent IMPs shortly after they leave the SecYEG translocon [25]. In addition, YidC functions independently of SecYEG in the integration of small IMPs such as the $F_oC$ subunit of ATP synthase and the bacteriophage coat protein M13 [26]. The mitochondrial YidC homologue Oxa1p from *S. cerevisiae* has been shown to directly interact with the ribosome [27, 28] but thus far ribosome binding has not been demonstrated for YidC, while the cytoplasmic domain of Oxa1p implied in ribosome binding is absent in YidC.

With the exception of some lactic acid bacteria, all completely sequenced bacterial genomes encode for the proteins SecD, SecF, and YajC. SecDFYajC forms a trimeric complex that is involved in protein translocation and associates with SecYEG [29, 30]. Two studies have indicated that SecDF might be both functionally and physically coupled to SecG [31, 32], but the exact function of SecDFYajC has remained elusive [33]. It has been proposed that SecDFYajC is involved in release of preproteins from the translocon, regulation of SecA cycling, and maintenance of the proton motive force. The latter proposal has been shown to be based on a polar effect of the growth conditions used with a SecDF depletion strain, rather than on the functional defects of the depletion of SecDF itself [34]. Further experiments are required to (dis)prove the other proposed functions of SecDF. In contrast to SecD and SecF, YajC is not required for cell viability. YajC alone has been shown to exist as a homooligomeric complex in the inner membrane of *E. coli* [35], but the functional importance of this complex is unknown.

## B. Evolutionary History of the *E. coli* Sec Machinery

Although the most intensively studied bacterial Sec machinery is that from *E. coli*, some characteristics of this system are not representative for the vast majority of bacteria. There are at least three components that distinguish the *E. coli* Sec machinery from that of other bacteria: SecB, SecM, and SecE. The tetrameric cytoplasmic protein SecB is a secretion specific chaperone that prevents intracellular aggregation of (pre)proteins [36]. SecB slows down the folding of preproteins by binding to their mature region [37], and it targets them to the extreme C-terminus of SecYEG-bound SecA [38]. Once translocation of the preprotein has been initiated, SecB is released from the translocon and able to start a new targeting cycle [11]. SecB is not essential for cell viability [39], but it is thought to be required for translocation of a subset of (pre)proteins [40]. Thus far, no clear amino acid motifs have been identified that render (pre)proteins SecB dependent [41], but it has been shown that SecB-binding sites are enriched in aromatic and basic residues [42].

The second component that distinguishes *E. coli* from most other bacteria is SecM, a small regulatory protein (formerly known as gene X) that is encoded directly upstream of SecA [43]. Under secretion-deficient conditions, SecM induces a pause in translation of the *secM–secA* messenger RNA by means of an arrest sequence in its C-terminus [44] that is sensed by the interior of the ribosome [45]. This results in prolonged exposure of the SecA ribosome-binding site and consequently an upregulation of the amount of cellular SecA. In addition, SecM is involved in localizing the expression of SecA to the vicinity of SecYEG [46]. SecM contains a signal sequence at its N-terminus, and thus the ribosome carrying a *secM–secA* messenger and the arrested nascent chain is targeted to the translocon. The SecA molecules that are subsequently synthesized in the vicinity of SecYEG are more active in protein translocation than SecA molecules that are synthesized without a functional *secM* gene in *cis* [46]. This SecA population possibly corresponds to the ''membrane integral'' form of SecA [47, 48]. SecM is not required for cell viability provided that sufficient SecA is supplied *in trans* [43].

SecE is the third component that distinguishes *E. coli* from many other bacteria; *E. coli* SecE consists of three transmembrane segments (TMSs), whereas most of its homologues are single spanning membrane proteins [49]. The additional two TMSs might specifically facilitate protein translocation at low temperatures, since *E. coli* cells containing a variant of SecE lacking these two TMSs are cold sensitive for growth [50].

An extensive genome analysis has revealed that SecB, SecM, and SecE with three TMSs are not unique to *E. coli* as they are present in several other proteobacteria, but not in any other bacterial divisions [51]. It is tempting to speculate that an optimized Sec machinery could be particularly beneficial to the frequently pathogenic proteobacteria, but it should be noted that the microbial genome-sequencing projects are strongly biased toward pathogenic organisms in general. Interestingly, the genomic distribution of SecB, SecM, and SecE with three TMSs reveals a part of the evolutionary history of the *E. coli* Sec machinery. By combining the genomic distribution with the phylogenetic relationships between the proteobacterial subdivisions in which each component is present (Figure 2.1), it was revealed that the Sec machinery has most likely evolved in the following successive steps: within the proteobacteria, the canonical Sec machinery (containing only SecYEG, SecA, SecDFYajC, and YidC) was first supplemented with SecB, then SecE was extended with two TMSs, and finally SecM was introduced. Hence, the *E. coli* Sec machinery represents the end product of a stepwise evolutionary process. Intermediate compositions with only SecB or SecB in combination with a three TMS-containing SecE are also observed, but neither the extended SecE nor SecM is ever observed without SecB, and SecM is

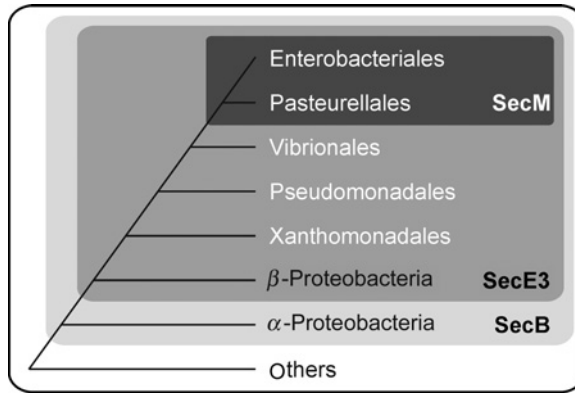



FIG. 2.1. Genomic distribution of accessory features of the Sec machinery in proteobacteria in combination with bacterial phylogeny. The distribution suggests that the Sec machinery has evolved in a stepwise fashion by sequentially acquiring SecB, the SecE extension, and SecM [51].

never observed without extended SecE. It has been proposed that both SecE with three TMSs and SecM could specifically improve SecB-dependent protein translocation by maximizing the amount of SecYEG-bound SecA that forms the receptor for preprotein–SecB complexes. This can be accomplished in two ways: (1) by increasing the affinity of SecA for SecYEG (via SecE) or (2) by carefully regulating and localizing the expression of SecA (via SecM) [51]. Further biochemical studies are required to investigate the possible synergistic contribution of SecB, SecM, and extended SecE to protein translocation. What should be kept in mind is that the Sec machinery of the model organism *E. coli* is of much greater complexity than that of most other bacteria.

## C. SEC PARALOGUES

Noncanonical compositions of the Sec machinery-containing paralogues of one or more components are also observed in many bacteria. Several genomes of organisms belonging to the divisions Actinobacteria (e.g., *Mycobacterium tuberculosis*) [52] and Firmicutes (e.g., *Listeria monocytogenes* and *Streptococcus gordonii*) [53, 54] encode for paralogues of SecA, and few of those bacteria encode for paralogues of SecY, SecE, and/or SecG as well. The genomes of the proteobacteria *Gluconobacter oxydans* and *Francisella tularensis* encode for SecB paralogues [51]. The genomic distribution of these paralogues has not yet been investigated in an

evolutionary context, and SecA2 is the only paralogue that has been studied genetically. It has been shown in both *M. tuberculosis* [55] and in *L. monocytogenes* [56] that SecA2 is important for pathogenicity but not for viability. These observations have led to the speculation that the accessory Sec machinery components of these Gram-positive bacteria might be functional equivalents of the pathogenicity related Type II–IV secretion systems found in many Gram-negative bacteria [56]. The thus far identified SecA2-dependent substrates do not have any functional characteristics in common. However, several substrates contain an atypical signal sequence or become glycosylated before translocation [56–62]. Interestingly, some SecA2-dependent substrates do not contain a signal sequence at all [55, 56]. It will be of great interest to investigate these and other features that distinguish SecA2 and the other paralogues from the canonical Sec machinery.

## V. SecA Structure, Function, and Dynamics

### A. The Involvement of SecA in Cotranslational Protein Translocation

The motor protein SecA is one of the largest and most complex bacterial proteins. It consists of multiple domains and it interacts with nearly all the other components involved in protein translocation: (pre)proteins, SecYEG, SecB, nucleotides, the cytoplasmic membrane, and possibly the ribosome. Although co- and posttranslational translocation reactions are mostly studied as individual pathways in *S. cerevisiae* and *E. coli*, respectively, there are several indications that the two pathways overlap. Most IMPs are translocated cotranslationally, but several IMPs contain large extracytoplasmic domains that are translocated in a SecA-dependent manner [25, 63–66]. This implies that SecA and the ribosome can either bind to the translocon simultaneously or that they can bind alternating to the translocon. Although simultaneous binding of SecA and the ribosome to SecYEG is structurally difficult to envisage (see chapter 2.VI), it has been shown that ribosomes and SecA do not compete for binding to SecYEG [67]. In addition, it has been demonstrated that SecA has a low but intrinsic ribosome-binding capacity, either alone [68, 69] or in conjunction with SecYEG [67]. Interestingly, ATP hydrolysis by SecA appears to induce the release of the ribosome from the translocon [67]. In this context, it should be stressed that during translocation of a large extracytoplasmic domain of an IMP by SecA, the ribosome would remain tethered to the translocon via the nascent chain rather than being truly released. The latter

would favor rebinding of the ribosome to the translocon for cotranslational continuation of the translocation process. Taken together, the co- and posttranslational protein translocation pathways are likely to be intertwined. Therefore, *in vitro* membrane protein insertion studies with SecA-dependent membrane proteins of varying topologies are eagerly awaited to further unravel this intricate process. In particular, special attention should be paid to the role of YidC and SecDFyajC during membrane insertion of SecA-dependent IMPs.

## B. The Overall Mechanism of Posttranslational Protein Translocation

In contrast to its possible role in cotranslational protein translocation, the role of SecA in posttranslational translocation is understood in much more detail due to extensive biochemical studies with purified components. This has resulted in the following widely accepted working model (Figure 2.2): in SecB-containing organisms, the cycle of posttranslational translocation starts with binding of a (pre)protein–SecB complex to SecYEG-bound SecA [11], on which the preprotein is transferred to SecA [70]. In organisms lacking SecB, the preproteins either bind directly to SecYEG-bound SecA, or are targeted to the translocon via binding to cytoplasmic or lipid-bound SecA. The subsequent binding of ATP to SecYEG-bound SecA induces a conformational change that results in insertion of the signal sequence into the translocon, and release of SecB (if present). At the same time, SecA is thought to insert partially into the translocon [71], and around 2.5 kDa of the mature domain of the preprotein is translocated [72, 73]. ATP hydrolysis results in release of the (pre)protein

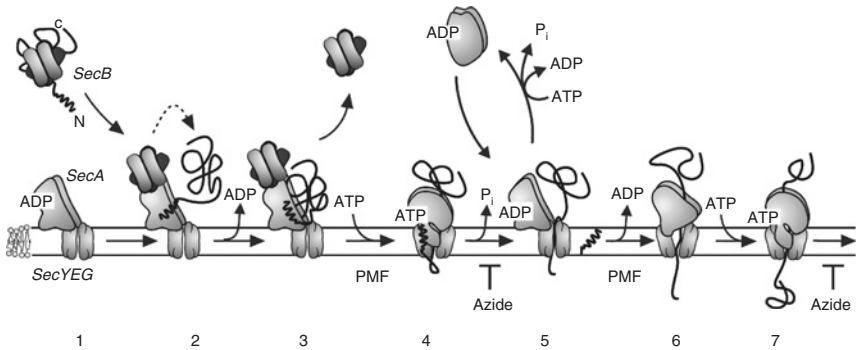



Fig. 2.2. Schematic representation of posttranslational protein translocation in *E. coli*. See text for details.

from SecA and deinsertion of SecA from the translocon, in a step that can be inhibited by the commonly used antibacterial compound azide [74]. Next, rebinding of SecA to the partially translocated polypeptide chain can drive the translocation of another 2.5 kDa of the mature preprotein domain [72, 73]. Depending on the length of the (pre)protein, multiple cycles of ATP binding and hydrolysis and SecA binding and release are required to completely translocate the substrate across the membrane.

## C. STRUCTURE OF THE SECA PROTOMER

The working model described above is still rather abstract, but our insight into the molecular details of the mechanism has become increasingly clear due to the availability of crystal structures from SecA [13, 17, 75], SecB [14, 15, 76], and an archaeal SecYEG homologue [16]. Three different crystal structures of SecA are available, two from *B. subtilis* and one from *M. tuberculosis*. The actual motor function of SecA, that is conversion of chemical energy into movement, is initiated by a ''DEAD motor'' core that is also present in DNA/RNA helicases [77]. The DEAD motor consists of two similarly folded domains that are referred to as nucleotide-binding folds (NBF1 and NBF2), each resembling the recombination protein RecA. At the interface of these two domains a single ATP molecule can be bound and hydrolyzed, which induces the conformational changes in SecA that ultimately results in the translocation of preproteins. SecA interacts with preproteins via the preprotein-binding domain (PBD, also referred to as ''preprotein cross-linking domain'' (PPXD) [78]) that is inserted into the amino acid sequence of NBF1 (Figure 2.3A), but forms a separate domain in the SecA structure [79, 13] (Figure 2.3B). The remainder of the SecA structure can be subdivided into four regions: the helical scaffold domain (HSD), the helical wing domain (HWD), the C-terminal linker (CTL), and the SecB-binding domain ''SecAc.'' The HSD forms a long scaffold to which NBF1, NBF2, the PBD, and the HWD are connected, the HWD is a loosely attached domain with unknown function, and the CTL forms the connection with SecAc at the extreme C-terminus [13] (Figure 2.3A and B).

### *1. Oligomeric State of SecA*

In order to understand the working mechanism of any protein on a molecular level, it is not only essential to know its structure and the exact location of the interaction sites for all its ligands but also to elucidate the functional oligomeric state of the protein itself. The oligomeric state of both SecA and SecYEG during protein translocation has become a controversial

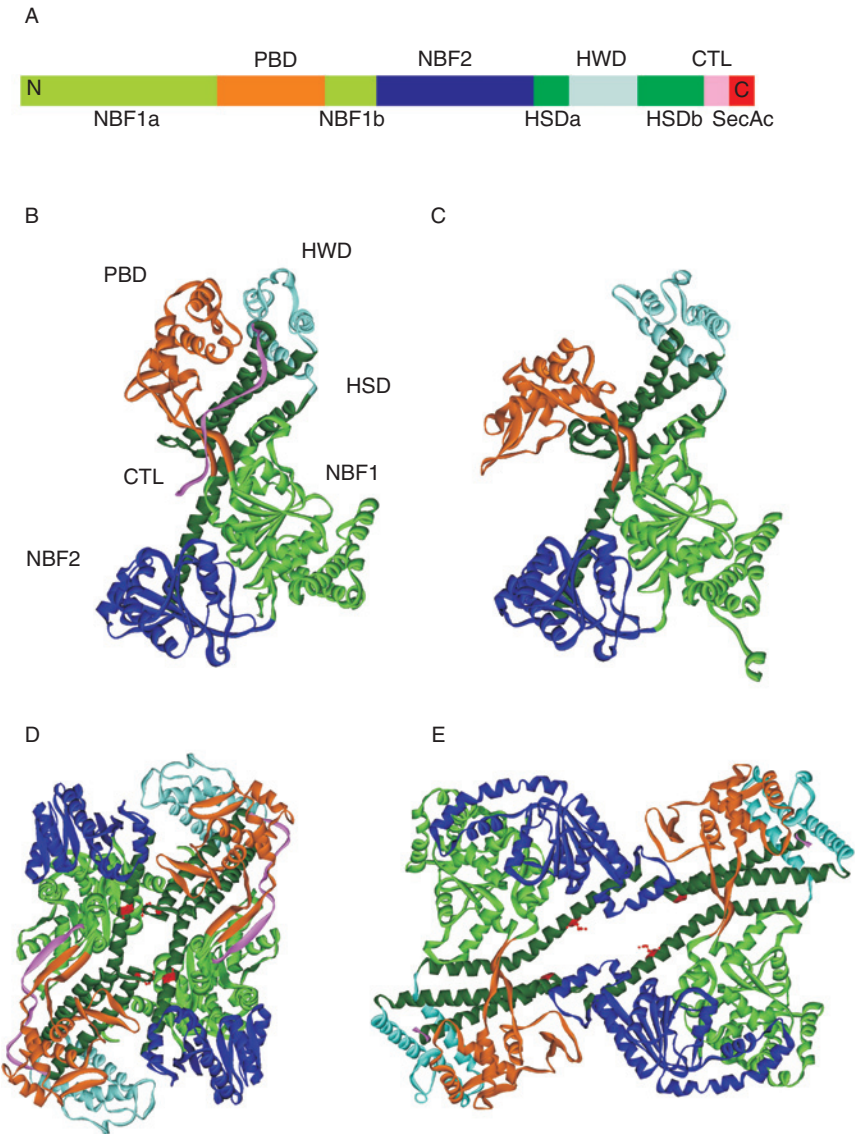



FIG. 2.3. Structure of SecA. (A) Schematic overview of the domain structure of SecA. NBF: nucleotide-binding fold; PBD: preprotein-binding domain; HSD: helical scaffold domain; HWD: helical wing domain; CTL: C-terminal linker; SecAc: SecB-binding motif. (B) Crystal structure of SecA protomer from *B. subtilis* with individual domains colored as in (A) [13]. (C) Crystal structure of SecA from *B. subtilis* in an open conformation, possibly representing the (pre)protein-bound state [75]. The conformational changes with respect to the structure

topic, and the complexity of the matter is schematically depicted in Figure 2.4. In an attempt to enlighten both discussions, we will address the topics individually, starting with SecA. For clarity, we have grouped the experimental data according to the following three subquestions:

I. What is the oligomeric state of soluble SecA?
II. What is the oligomeric state of SecYEG-bound SecA?
III. What is the oligomeric state of translocation-engaged SecA?

### 2. *The Oligomeric State of Soluble SecA*

It has been shown with various techniques that purified SecA exists in a dynamic equilibrium between a monomeric and a dimeric form, and the dissociation constant ($K_D$) has been estimated to be around 0.1 μM under physiological conditions [80]. The cellular concentration of SecA is ~8 μM [81], and thus SecA is expected to be largely dimeric *in vivo*. Higher order SecA oligomers have also been reported, but only under nonphysiological conditions or with truncated SecA mutants [17, 82]. Three reports have shown that translocation ligands can induce monomerization of SecA dimers, which raises the question whether the cellular predominant SecA dimer is also the functional state. Fluorescence- and cross-linking studies with purified SecA have shown that the monomer–dimer equilibrium can be shifted toward the monomer by the addition of certain lipids or detergents [83, 84], or signal peptides [83, 85], although a different view has been

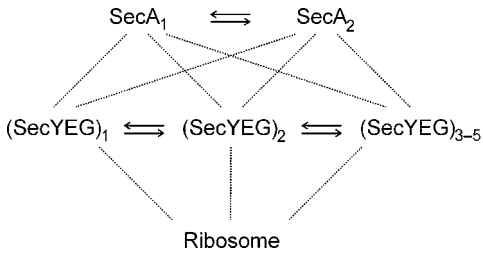



FIG. 2.4. Schematic overview depicting the complexity of the debate concerning the oligomeric states of SecA and SecYEG. The experimentally demonstrated equilibria between the different oligomeric states are indicated by arrows, and all the possible interactions are indicated by dashed lines. See text for details.

depicted in (B) are indicated by arrows. (D) Crystal structure of dimeric SecA from *B. subtilis* that most likely represents the physiologically active dimer [13]. The two intradimeric HSD–HSD contacts that are maintained during protein translocation are depicted in red [95]. (E) Crystal structure of *M. tuberculosis* SecA. (See color plate section in the back of the book.)

published, suggesting that signal peptides induce oligomerization of SecA [84]. Lipid-bound SecA has been shown to exist mainly in a dimeric form that can be dissociated on binding of nucleotides [86]. Although all these studies underscore the dynamic and sensitive nature of the SecA monomer–dimer equilibrium, it remains questionable whether any of these observed changes in oligomeric state are functionally relevant since no SecYEG nor (pre)proteins were present in these studies.

### *3. The Oligomeric State of SecYEG-Bound SecA*

The oligomeric state of SecA while bound to SecYEG detergent solution has been addressed by native gel electrophoresis and gel filtration [87–89]. It was shown that both monomeric and dimeric SecA can bind to SecYEG, provided that SecYEG is stabilized in a dimeric form either by covalent linkage [88] or by an antibody [89]. Unstabilized SecYEG in detergent only retains monomeric SecA after a preprotein has been trapped inside the channel before solubilization of the membrane [88]. These results should be interpreted carefully, however, since the monomer–dimer equilibrium of SecA has been shown to be highly sensitive to detergents [83].

The oligomeric state of SecA bound to membrane-embedded SecYEG has been addressed by chemical cross-linking [90] and surface plasmon resonance (SPR) [91]. Dimeric SecA can be detected after binding to urea-treated inverted membrane vesicles (IMVs) [90], but could not be detected with SecYEG-containing proteoliposomes [83]. The concentration of SecA added in the latter experiment however was far below physiological (5 nM vs 8 μM), and thus the results obtained with the IMVs appear to be more reliable. Chemical cross-linking of the population of SecA that copurifies with IMVs revealed mainly SecA monomers [90], while the fraction of SDS-resistant dimers dramatically increases on overexpression of SecYEG [91]. SPR measurements also suggest that SecA is dimeric while it is bound to membrane-embedded SecYEG, since wild-type SecA binds to SecYEG-overexpressing IMVs similarly to a covalently cross-linked SecA dimer [91]. Taken together, these data indicate that both monomeric and dimeric SecA can bind to SecYEG.

### *4. The Oligomeric State of Translocation-Engaged SecA*

Activity assays are obviously the most relevant experiments to assess the oligomeric state of SecA during protein translocation. In order to investigate the functional requirement of dimeric SecA, several studies have characterized SecA mutants with disturbed dimerization properties. Removal of the N-terminal eight amino acids of SecA does not influence its oligomeric state

[92], but SecA has been reported to be predominantly monomeric when the first 11 amino acids are removed [90, 93, 94]. Alternatively, monomeric SecA can be obtained by mutating 6 residues of a SecA truncate that lacks 70 residues from its extreme C-terminus [83]. It should be noted that these monomeric SecA mutants are not incapable of dimerization per se, as the mutations have shifted the monomer–dimer equilibrium in solution substantially toward the monomeric state [93]. In the two assays that measure SecA activity, for example the *in vitro* preprotein translocation assay and the precursor-stimulated SecA ATPase assay, all monomeric SecA mutants show either a very low activity or no activity at all. Although the low residual *in vitro* activity has been interpreted as being significant in some reports [83, 94], it seems more likely that the residual activity is caused by a small fraction of SecA dimers that can still be formed or by traces of copurified wild-type SecA from the expression host.

Activity assays with covalently dimerized SecA have yielded varying results. SecA dimers cross-linked via endogenous cysteines located in the SecB-binding domain (SecAc) [91] or via a pair of engineered cysteines in the HSD ($Arg^{637}$ and $Gln^{801}$) [95] were shown to be nearly fully active in protein translocation and preprotein-stimulated SecA ATPase activity. Although these observations alone do not directly imply that SecA functions as a dimer, it does show that monomerization is not required for functionality as proposed earlier [83].

Perhaps the most convincing experiment that assesses the functional oligomeric state of SecA-involved heterodimers of active and inactive SecA monomers [96]. If SecA would function as a monomer, these heterodimers are expected to have half the activity of wild-type SecA. However, it was observed that these heterodimers are completely inactive, strongly suggesting that SecA is functional as a dimer.

### 5. *Summary of Oligomeric States SecA*

Taken together from our point of view, the experimental data showing that SecA dimers dissociate on binding of translocation ligands are not necessarily related to protein translocation, since they might simply reflect the sensitive nature of the monomer–dimer equilibrium. The data supporting the proposal that SecA functions as a monomer are in our opinion either; obtained under conditions too distant from physiological; explainable by a conformational change of SecA, or misinterpreted. On the other hand, the experimental data supporting the SecA dimer as a functional unit are more convincing and more abundant. Furthermore, there are no experimental data disproving a functional SecA dimer, whereas *in vivo* and *in vitro* experiments in different laboratories demonstrate that monomeric

SecA variants are inactive. Finally, it has been shown that SecB targets preproteins to dimeric SecA, and that this targeting greatly stimulates the efficiency of protein translocation [38]. Combined with the notion that cellular SecA is predominantly dimeric, we assume that SecA functions as a dimer in posttranslational protein translocation at SecYEG.

We speculate that the physiological relevance of the binding of monomeric SecA to SecYEG and the sensitive nature of the monomer–dimer equilibrium could be related to (pre)protein targeting to SecYEG. As mentioned above, in organisms lacking SecB, (pre)proteins might first bind to cytoplasmic or lipid-bound SecA, and subsequently transferred to the translocon. If one SecA protomer would remain permanently bound to SecYEG, the dimerization of SecA could play a role in the initiation of translocation via this SecB-independent targeting process.

### D. Structure of the Functional SecA Dimer

With our current insight that SecA functions as a dimer, the next question is at which side of a SecA protomer the intradimeric interactions take place. Two interactions observed in various crystal structures have been proposed to represent a physiological dimer interface [13, 17]. The overall arrangement of both of these SecA dimers is very similar; the two elongated SecA monomers are arranged side-by-side in an antiparallel fashion (Figure 2.3D and E). This antiparallel arrangement is supported by fluorescence resonance energy transfer (FRET) [97] and cross-linking studies [94, 95]. The difference between both dimers lies in the SecA surface that contacts the neighboring protomer. The proposed *B. subtilis* dimer is relatively compact and the dimer interface comprises a large surface (5442 $Å^2$) [13], whereas the *M. tuberculosis* dimer is relatively flat, comprises a smaller surface (2822 $Å^2$), and contains a cavity at the dimer interface [17]. One dimer arrangement can be converted into the other by rotating each protomer $\sim 75^\circ$ around its long axis. Although it is conceivable that such rotations could play a role in the cycle of SecA-driven protein translocation, the observation that a SecA dimer that is fixed in the *B. subtilis* arrangement (Figure 2.3D) still supports efficient protein translocation [95] suggests that at least the *B. subtilis* dimer is part of the conformational cycle of SecA. Thus, it can be concluded that the HSDs of two SecA protomers can be considered as a single scaffold domain in the SecA dimer, and that none of the conformational changes that SecA undergoes during protein translocation is severely hampered by the intradimeric HSD–HSD cross-links. Whether the *M. tuberculosis* dimer arrangement (Figure 2.3E) also represents a functional intermediate remains to be established.

## E. CONFORMATIONAL CHANGES WITHIN SECA

Several regions in SecA have been shown to be dynamic [13, 98–105], but detailed structural information is only available on two conformational changes: one that can be inferred from SecA's similarity to helicases and another that has been observed directly with X-ray crystallography [75]. As mentioned previously, the DEAD-motor core of SecA (NBF1 and NBF2) is homologous to that of SFI and SFII helicases, and therefore the nucleotide-induced conformational changes are assumed to be similar in all three protein families. SecA has been crystallized with bound ADP and in the nucleotide free state, but these structures differ only slightly in the orientation of side chains that are involved in nucleotide binding. Unfortunately, attempts to crystallize SecA in the functionally important ATP-bound state have failed thus far. In addition to conformations that are very similar to those of nucleotide free and ADP-bound SecA, the helicase DEAD motors have been crystallized in two substantially different conformations. First, the SFII helicase MJ0669 has been crystallized without nucleotides in an open conformation in which the two NBFs are separated from each other by a large cleft [106]. Second, the SFI helicase PcrA has been crystallized in the ATP-bound state in which the two NBFs have undergone an $\sim 10^{\circ}$ rotation relative to each other compared to the ADP-bound state [107]. All three distinct conformations as observed in different DEAD motors (open, closed, and closed-rotated) are assumed to underlie the ATPase cycle of SecA as well. Given the observation that a SecA dimer in which the two HSDs are cross-linked is still active, the relative reorientations of NBF1 and NBF2 that are required for ATP binding and hydrolysis are apparently not influenced by these disulfide-bonded cross-links. When the mobility of NBF1 is restricted by a disulfide cross-link to the HWD of the neighboring protomer however, the SecA dimer is inactive [95].

The conformational change of SecA that has been visualized by X-ray crystallography does not involve the DEAD-motor or nucleotide, and it takes place in the opposite end of a SecA protomer [75]. *B. subtilis* SecA has been crystallized in two different conformations, and a comparison of both conformations reveals the following movements in a protomer: the HSD and HWD undergo a small rotation, and the PBD undergoes a large ($\sim 60^{\circ}$) rotation combined with a rigid body translation away from the HSD and HWD (Figure 2.3B and C). This results in opening of a groove at the PBD–HSD/HWD interface (Figure 2.3C) that has been proposed to form the actual preprotein-binding site since its physicochemical characteristics are similar to that of peptide-binding sites from other proteins with broad substrate specificities. Assuming that this conformation of SecA represents a (pre)protein-bound state and knowing that *B. subtilis* does not contain a SecB protein

it could represent either a SecYEG-bound form, a lipid-bound form, or a soluble form. In the latter two cases, it might represent the earlier proposed (monomeric) form of SecA that was suggested to be involved in SecB-independent targeting of (pre)proteins to a SecYEG-bound protomer. As at present it is unclear whether the observed conformational changes can take place in the *B. subtilis* dimer arrangement, the conformation in the crystal structure could also represent (one of) the SecYEG-bound SecA protomer(s) after receiving a (pre)protein. The location of the CTL that connects the SecB-binding domain SecAc to the HWD suggests how binding of a SecB-(pre)protein complex could be mechanistically coupled to the conformational change in SecA (see chapter 2.V.F).

## F. SecA–SecB Interaction

The interaction between SecA and SecB has been investigated in great detail. Since an excellent review on the SecA–SecB interaction has appeared [108], we will only discuss the most important findings and a possible relation to conformational changes in SecA. It has been shown that the extreme C-terminus of SecA (SecAc) contains a dedicated SecB-binding site that is formed by a small cysteine-rich domain that chelates a zinc ion [109]. This highly conserved domain is also found in organisms lacking SecB, which might be related to the fact that the C-terminus is also involved in lipid binding [110]. The SecAc domain is not resolved in any of the available SecA crystal structures, but its structure has been determined in isolation by NMR [111, 112] and in complex with *Haemophilus influenzae* SecB by X-ray crystallography [15]. The latter structure revealed that two SecAc domains are bound to opposite sides of one SecB tetramer, on a surface that was previously shown to be crucial for SecB-binding to SecA [70, 113]. The SecAc domain is stabilized by the zinc ion that is coordinated by three cysteines and one histidine, explaining why SecA mutants in which these residues are either mutated [114] or cross-linked [91] are unable to support SecB-dependent protein translocation.

The approximate position of the SecB tetramer bound to SecA in the *B. subtilis* dimer arrangement has been estimated by docking of the SecB–SecAc complex onto the SecA structure [108]. It seems likely however that on binding of a preprotein–SecB complex to SecA, the transfer of the (pre)protein requires (or induces) a substantial conformational change in SecA [70]. This conformational change possibly corresponds to the one that is observed by X-ray crystallography [75]. Binding of SecB to the highly mobile SecAc domain could displace the CTL that connects SecAc to the HWD. Since the CTL is part of the PBD-hinge region in the closed conformation of SecA and it meanders partially underneath the PBD

(Figure 2.3B), this displacement could directly induce the observed rigid body movement of the PBD that results in opening of the proposed (pre) protein-binding groove (Figure 2.3C). Furthermore, CTL displacement could be directly responsible for the small rotation of the HWD/HSD that coincides with opening of the groove. Although *B. subtilis* does not contain SecB, it has been shown that *E. coli* SecA undergoes a similar conformational change [75]. In organisms lacking SecB, displacement of the CTL is expected to be induced by an alternative mechanism. This could involve the interaction of SecAc with lipids [110] or binding of SecA to SecYEG [115].

### G. SECA–MEMBRANE INTERACTION

A detailed understanding of SecA binding to the membrane is fundamental for understanding the molecular mechanism of SecA-driven protein translocation. However, whereas binding of SecA to *E. coli* membranes has been studied extensively, surprisingly little is known about the region(s) of SecA that interact(s) with the membrane. The lipid-binding region of SecA has been localized to its C-terminal 70 amino acids [110], but the SecYEG-binding region of SecA has not been identified in detail. Far western experiments using SecA fragments mapped the SecYEG-binding region to the N-terminal part of the SecA protomer, comprising both NBFs and the PBD [116]. Moreover, binding experiments with SecA fragments have demonstrated that the same N-terminal region of SecA comprises the high-affinity SecYEG-binding site, whereas the remaining C-terminal one-third of SecA does not bind to SecYEG [116]. However, the exact SecYEG interaction sites within the N-terminal region have not been determined yet. The relatively new technique of cysteine-directed cross-linking in combination with mass spectrometry appears to be the most suitable biochemical approach to identify the exact regions in SecA that interact with SecYEG. In addition, medium- and high-resolution structural studies on SecYEG–SecA complexes will contribute to answering this critical question.

## VI. SecYEG Structure, Function, and Dynamics

### A. STRUCTURE OF THE SECYEG PROTOMER

The structure–function relationship of the translocon has been extensively studied in *E. coli* and *S. cerevisiae*. The recently solved high-resolution translocon structure from the archaeon *Methanococcus jannaschii* [16] was a major breakthrough in the field. Despite the fact that archaeal translocon subunits are more similar to eukaryotic than to bacterial ones [3], they are

commonly named after the bacterial subunits. Since no significant sequence similarity can be detected between SecG and its archaeal counterpart Sec(61)β [117], the eukaryotic nomenclature is applied to the latter, resulting in the hybrid term SecYEβ. In agreement with its universal conservation, the overall structure of *M. jannaschii* SecYEβ is nearly identical to that of *E. coli* SecYEG [118]. The two complexes differ only slightly in conformation [119], and the *E. coli* translocon contains three additional TMSs compared to that from *M. jannaschii*: two from SecE (section IV) and one from SecG. The center of the complex is formed by SecY, whereas SecE and SecG are located at the periphery (Figure 2.5A and B). The structure of

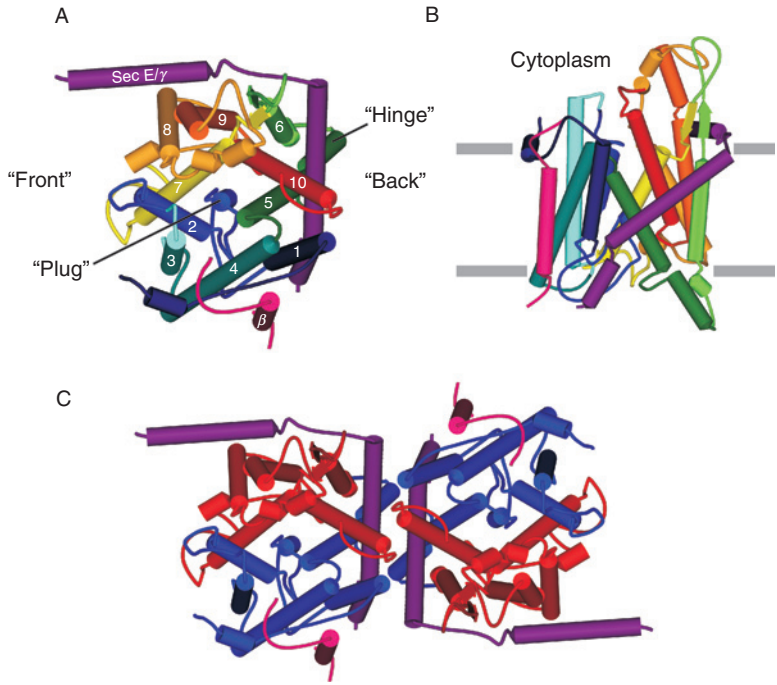



FIG. 2.5. Structure of SecYEβ from *M. jannaschii* [16]. (A) Cytoplasmic view showing the arrangement of transmembrane segments in different colors. SecE is depicted in purple, Secβ in pink. Sides referred to as ''front'' and ''back'' are indicated. (B) View from within the plane of the membrane showing the two cytoplasmic loops that extend into the cytoplasm and have been shown to interact with the ribosome and SecA: C4 and C5, connecting TMS6 with TMS7 and TMS8 with TMS9, respectively. (C) Back-to-back dimer arrangement of SecYEβ protomers as observed for *E. coli* SecYEG in two-dimensional crystals [118]. The N-terminal halves of SecY are depicted in blue, the C-terminal halves in red, and SecE and Secβ in purple and pink, respectively. (See color plate section in the back of the book.)

*M. jannaschii* SecYEβ consists of two distinct domains that are similarly folded. Each domain is composed of a bundle of five TMSs, formed by the N- or C-terminal half of the SecY sequence, respectively. The two halves of SecY are held together by SecE: the conserved TMS of SecE crosses the membrane diagonally [120], and contacts both SecY halves at the same side where they are connected by the extracytoplasmic loop between TMS5 and TMS6. This side of the SecYEG protomer is referred to as the ''back.'' The amphipathic cytoplasmic helix of SecE [121] runs parallel to the membrane surface along the C-terminal half of SecY (Figure 2.5A). Two of the cytoplasmic loops of SecY protrude far into the cytoplasm: the C4 loop connecting TMS6 with TMS7, and the C5 loop connecting TMS8 with TMS9 (Figure 2.5B). The extracytoplasmic loops on average are considerably shorter, and two of those fold back into the membrane region: the E4 loop connecting TMS7 with TMS8, and the E1 loop connecting TMS1 with TMS2. The latter is highly conserved, folds back between the two SecY halves, and is referred to as the ''plug'' domain [16].

At first sight, there is no obvious region in the channel that is large enough to allow passage of unfolded proteins. For this reason, it has been concluded that the structure represents the closed conformation of SecYEβ. However, on the basis of two domain structure of the channel and the observation that signal sequences of (pre)proteins can be cross-linked to TMS2 and TMS7 [122–124] at the domain interface, it was proposed that insertion of the signal sequence between TM2 and TM7 results in separation of the two halves of SecY and displacement of the ''plug'' that blocks the proposed pore from the extracellular side and that the substrates pass through the center of the channel [16]. Molecular dynamics simulations have revealed that the opening that is created by this mechanism is indeed large enough to allow passage of unfolded and even α-helical proteins [125]. In the opened state, nascent IMPs (and signal sequences) could leave the translocon laterally toward the lipid bilayer via the TMS2–TMS7 interface. The possible mechanisms by which SecA or the ribosome could induce channel opening will be discussed below, but first we will address the oligomeric state(s) of the translocon.

### *1. Oligomeric States of SecYEG*

As outlined above for SecA, knowledge of the functional oligomeric state of a protein is of fundamental importance for understanding its mechanism of action. Also the oligomeric state of SecYEG is heavily debated (Figure 2.4). In an attempt to enlighten this discussion, we will give an overview of the relevant experimental data. For clarity, we have

subdivided the assessment of the oligomeric state of SecYEG into three subquestions:

I. What is the oligomeric state of SecYEG in the absence of ligands?
II. What is the oligomeric state of SecYEG with bound SecA?
III. What is the oligomeric state of SecYEG with a bound ribosome?

### 2. *Oligomeric State of SecYEG in Absence of Ligands*

The oligomeric state of SecYEG in the absence of ligands has been addressed with several cross-linking studies and fluorescence resonance energy transfer (FRET). All these studies indicate that at least two copies of SecY [126–128], SecE [129, 130], and SecG [131] are present in a single complex. However, whether such an oligomeric complex contains two or more copies of each subunit can not be distinguished. More accurate information on the oligomeric state of purified SecYEG has been obtained in detergent solution by density centrifugation [132], analytical ultracentrifugation [133], gel filtration [89], native gel electrophoresis [87], and negative stain EM [132, 134, 135]. Several of these studies indicate that SecYEG exists in a dynamic equilibrium between monomers, dimers, and larger oligomers. The latter group includes presumed trimers, tetramers, and pentamers. Similar results were obtained with SecYEG reconstituted into lipid bilayers [130].

The observation of trimeric/tetrameric purified SecYEG complexes, per se, does not necessarily imply that these oligomeric states are also functionally relevant. Concerning this aspect three critical comments should be given. First, most of the experimental conditions that addressed the oligomeric state of SecYEG involve high concentrations of (overexpressed) SecYEG, and these might lead to nonphysiological distributions of the oligomeric states [136]. Second, the removal of SecYEG from a potential ''supercomplex'' with SecDFYajC and/or YidC in the membrane [29, 33, 137] might expose surfaces on SecYEG that in absence of these subunits could form an interaction site for self-association. Third and most importantly, the oligomeric state of SecYEG during protein translocation, that is with bound ligands, might differ from that in a ''resting'' state.

### 3. *Oligomeric State of SecYEG with Bound SecA*

SecA has been shown to bind to both dimeric [88, 89] and tetrameric SecYEG [130, 134], but not to SecYEG monomers [88, 89]. Binding of SecA induces a shift in the SecYEG equilibrium from monomers to dimers and tetramers, both in detergent solution [89] and in lipid bilayers [130]. In addition (membrane insertion of) SecA has been shown to increase the

amount of SecYEG dimers and proposed tetramers at the expense of SecYEG monomers [130, 134]. Constitutive SecYEG dimers that were created by covalent linkage [88] (N. Nouwen, unpublished data) or via disulfide cross-linking [127] were shown to be active in posttranslational protein translocation. Taken together, all these data indicate that in contrast to an earlier proposal [138] SecYEG functions in posttranslational translocation as an oligomeric complex. The exact oligomeric state however is difficult to assess, as pro- and contraarguments can be given for both dimers and higher order oligomers.

### *4. Oligomeric State of SecYEG with a Bound Ribosome*

The oligomeric state of the translocon is not necessarily the same during the post- and cotranslational translocation modes. The oligomeric state of the translocon during cotranslational translocation has been studied in both bacteria and eukarya, mainly by EM. Early EM studies of rough ER membranes revealed the existence of large ringlike particles that were estimated to contain three to four translocons [135]. Importantly, the formation of these particles from purified and membrane-reconstituted translocons was induced by the addition of ribosomes. Several subsequent cryo-EM studies on eukaryotic ribosome-bound translocons revealed that irrespective of the presence of an arrested nascent chain, similarly sized particles bind to ribosomes [18, 20, 21, 139, 140, 141]. Recently, however, a cryo-EM reconstruction of an *E. coli* ribosome-bound translocon was presented that was estimated to consist of only two SecYEG protomers, despite the fact that the overall size of this translocon is similar to the other reconstructions [19]. Given the universal conservation of cotranslational protein translocation and the observation that the ribosome–translocon interaction is conserved across the three domains of life [142], it seems unlikely that this difference reflects a property that distinguishes the bacterial translocon from its eukaryotic counterparts. A conclusive assessment of the oligomeric state of the ribosome-bound translocon is limited by the medium resolution of the currently available cryo-EM structures.

### *5. Summary Oligomeric States SecYEG*

Taken together, the oligomeric state of SecYEG during both co- and posttranslational protein translocation is at least dimeric, but the exact number of protomers constituting an active translocon remains controversial. Biochemical data assessing the oligomeric state of SecYEG during cotranslational translocation in particular and higher resolution three-dimensional structures of ribosome-bound translocons are eagerly awaited to resolve this critical issue.

## B. Arrangement of SecYEG Protomers Within an Oligomeric Assembly

Since the oligomeric state of SecYEG during both co- and posttranslational translocation is at least dimeric, it is relevant to assess the arrangement of SecYEG protomers within a dimeric assembly. By fitting the high-resolution structure of *M. jannaschii* SecYEβ into a previously solved three-dimensional reconstruction of *E. coli* SecYEG based on two-dimensional crystals [118], it was revealed that the conserved TMS of SecE is located at the dimer interface. Several cross-linking studies showed a similar localization of SecE in SecYEG complexes within *E. coli* inner membrane vesicles [129]. Importantly, several covalent linkages of constitutive SecY dimers that do not interfere with activity [88, 127] span the same dimer interface, suggesting that this so-called back-to-back arrangement could represent a physiological SecYEG dimer. Tetrameric assemblies of SecYEG have been proposed to consist of two back-to-back dimers arranged side-by-side (a dimer of dimers) [140], such that SecG and the amphipathic helix of SecE are located at the interface of the two dimers. However, this specific tetrameric arrangement is not supported by structural data, while SecG-dependent tetramerization is only supported by scarce biochemical evidence [87].

On the basis of cryo-EM reconstruction of ribosome-bound *E. coli* SecYEG, a radically different dimer arrangement of SecYEG protomers was proposed [19]. For generation of stable RNCs, the SecM translation arrest sequence was used and the complex that was isolated consists of the 70S ribosome (50S and 30S subunit) carrying a nascent single-spanning membrane protein, mRNA, three tRNAs, and two translocons. One of the translocons is bound to the arrested nascent chain at the polypeptide exit tunnel as observed in previous studies, but the other is bound to the mRNA via an interaction that is most likely nonphysiological. On the basis of normal mode flexible fitting (NMFF) of SecYEG into the observed electron densities, it was proposed that the two translocons represent SecYEG dimers in a front-to-front arrangement in an open and a closed conformation, respectively (Figure 2.6B and A). Importantly, these analyses suggested that the conformational change underlying opening of the channel indeed involves separation of the two SecY halves. Prominent electron density that most likely corresponds to the arrested nascent chain was observed at the TMS2–TMS7 interface of the two neighboring SecY molecules (black cross in Figure 2.6B), rather than at the TMS2–TMS7 interface of a single SecY. This led the authors to propose that after being inserted into a single SecYEG protomer at the interface of the two SecY halves, nascent membrane proteins leave the translocon laterally via the

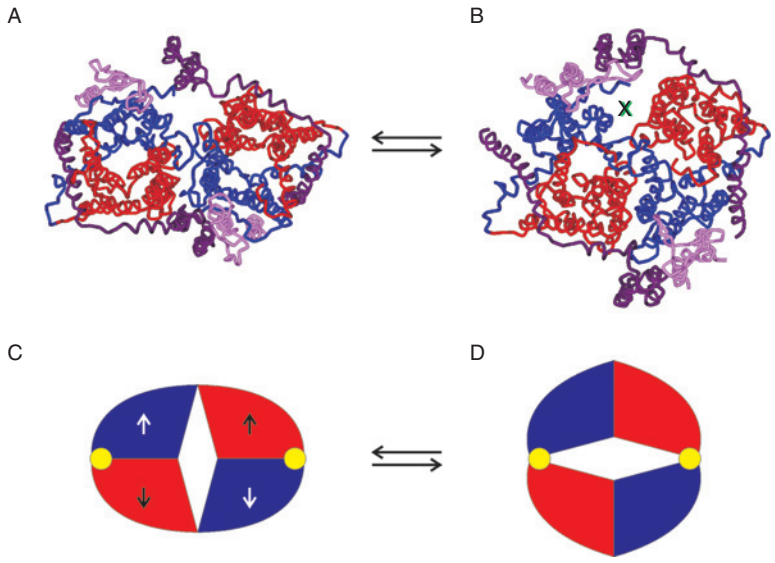



FIG. 2.6. Front-to-front dimer arrangements of *E. coli* SecYEG [19]. (A) Closed conformation of the front-to-front dimer, nonphysiologically bound to mRNA in the cryo-EM structure. (B) Open conformation of the front-to-front dimer bound to the arrested nascent chain at the ribosomal exit tunnel. The black cross indicates the position of the electron density that possibly corresponds to the arrested nascent chain. In (A) and (B), the N-terminal halves of SecY are depicted in blue, the C-terminal halves in red, SecE in pink, and SecG in green. (C) and (D) Schematic representation of the proposed ribosome/SecA-induced opening mechanism. A simultaneous interaction of the ribosome or SecA with the N-terminal (blue) and C-terminal (red) domain of one or two SecY molecules could induce opening of the translocon via outward directed forces. The proposed hinge region (loop E3 connecting TMS5 and TMS6) is represented by yellow circles, the proposed outward directed forces are indicated by arrows. The large clefts within both states of the translocon are merely for illustrative purposes. (See color plate section in the back of the book.)

interface of two SecY molecules. Furthermore, the front-to-front arrangement will allow the formation of a large consolidated pore that could be required for hairpin insertion of (pre)proteins and/or translocation of substrates containing bulky side chains or internal disulfide bonds [143, 144]. Although other cryo-EM studies consistently indicated translocon oligomeric states of a higher order than dimers and a front-to-front arrangement of protomers was unanticipated, the proposed model provides many explanations for previously obtained biochemical results. Future biochemical and structural studies are required to experimentally validate the proposed front-to-front model.

### C. Induction of Conformational Changes in SecYEG

Assuming that the proposed open conformation of dimeric SecYEG represents a physiologically active translocon, the question is how the ribosome or SecA can induce opening of the channel. Interestingly, the ribosome and SecA interact with similar regions of the translocon, suggesting that they might share a common opening mechanism. The ribosome interacts with the translocon via three distinct connections. In agreement with biochemical studies [145, 146], two connections are similarly formed by the pairs of long cytoplasmic loops of SecY (C4 and C5, connecting TMS6 with TMS7 and TMS8 with TMS9, respectively, black arrows in Figure 2.6C). The third connection is mediated by the cytoplasmic loop of SecG and the N-terminus two transmembrane segments of SecE (one of the white arrows in Figure 2.6C). SecA has been shown to interact with the C5 loop of SecY as well [158], with SecG [147], and with the interface between TMS4 and C3 of SecY (EvdS *et al.* submitted for publication) that is in direct contact with SecG [127]. Importantly, the two regions of interaction are located in different domains of a single SecYEG protomer, and thus separation of the two SecY domains could be induced by a simultaneous interaction with both of them (Figure 2.6C and D). In the front-to-front arrangement, separation of the two SecY domains mainly takes place at the dimer interface, and thus opening of a single protomer will be directly transmitted to the neighboring protomer.

It should be noted, however, that the features that mediate the third ribosome–translocon connection (SecG/Secβ and the SecE extension) are not essential for viability or protein translocation [50, 148, 149]. Thus, ribosome-induced opening of the translocon might be primarily mediated by the two C4/C5 connections, while the third connection plays an auxiliary role. This would explain the mere stimulatory role of Secβ on posttranslational protein translocation [150]. The stimulatory role of SecG can be explained similarly, but the SecA-induced opening mechanism differs in at least one aspect from the ribosome-induced opening mechanism, that is the SecA ''membrane insertion'' cycle. The SecA interaction site in the N-terminal half of SecY (the TMS4–C3 interface) appears to be part of the region where SecA inserts at least partially into the translocon. SecG is in proximity of this region and might thus facilitate membrane cycling of SecA [151, 152]. It seems unlikely however that SecG completely inverts its membrane topology during protein translocation via SecYEG as proposed previously [153], as topologically fixed SecG has been shown to be equally active as wild-type SecG [154]. The different conformations of SecG that are observed *in vitro* most likely represent conformational changes within this highly dynamic region of the translocon.

### D. THE ROLE OF THE PLUG

In addition to separation of the two SecY domains, the opening mechanism of the translocon is thought to involve displacement of the ''plug'' domain formed by the E1 loop [16]. This proposal is based on the location of the plug domain at the extracellular end of the pore region in the closed conformation of the channel, and the observation that it has the potential to be cross-linked to the C-terminal region of SecE, located ~20 Å away [155]. The mobile nature of the plug domain has been confirmed by molecular dynamics simulations [125, 157], homology modeling [119], and a cross-linking approach [156]. In the latter study, it was shown that the plug domain is displaced during protein translocation, providing the first experimental evidence for its proposed function. An interesting observation that provides a novel hypothesis on the mechanism of SecA-induced opening of the translocon was recently made with a peptide scanning approach [158]. It was shown that SecA directly interacts with peptides derived from the plug domain, suggesting that displacement of the plug domain in bacteria might be directly induced by SecA.

## VII. Concluding Remarks

To summarize, our understanding of the molecular mechanism of protein translocation in bacteria has increased dramatically during the past few years, and long held schematic models are slowly beginning to take shape on a detailed structural level. However, a ''molecular movie'' of protein translocation is not expected in the near future because of the tremendous complexity of the process. New insights have to be provided by a combination of structural, biophysical, and biochemical studies. Considering the large amount of unresolved questions, research on the Sec machinery is expected to remain an exciting area in biology throughout the next decade.

### REFERENCES

1. Pohlschroder, M., Prinz, W.A., Hartmann, E., and Beckwith, J. (1997). Protein translocation in the three domains of life: Variations on a theme. *Cell* 91:563–566.
2. Osborne, A.R., Rapoport, T.A., and Van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. *Annu Rev Cell Dev Biol* 21:529–550.
3. Cao, T.B., and Saier, M.H. (2003). The general protein secretory pathway: Phylogenetic analyses leading to evolutionary conclusions. *Biochim Biophys Acta* 1609:115–125.
4. Ulbrandt, N.D., Newitt, J.A., and Bernstein, H.D. (1997). The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* 88:187–196.

5. Kalies, K.U., Gorlich, D., and Rapoport, T.A. (1994). Binding of ribosomes to the rough endoplasmic reticulum mediated by the Sec61p-complex. *J Cell Biol* 126:925–934.
6. Pohlschroder, M., Hartmann, E., Hand, N.J., Dilks, K., and Haddad, A. (2005). Diversity and evolution of protein translocation. *Annu Rev Microbiol* 59:91–111.
7. Rapiejko, P.J., and Gilmore, R. (1997). Empty site forms of the SRP54 and SR alpha GTPases mediate targeting of ribosome-nascent chain complexes to the endoplasmic reticulum. *Cell* 89:703–713.
8. Mason, N., Ciufo, L.F., and Brown, J.D. (2000). Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J* 19:4164–4174.
9. Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta* 1694:17–35.
10. Wild, K., Halic, M., Sinning, I., and Beckmann, R. (2004). SRP meets the ribosome. *Nat Struct Mol Biol* 11:1049–1053.
11. Hartl, F.U., Lecker, S., Schiebel, E., Hendrick, J.P., and Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* 63:269–279.
12. Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc Natl Acad Sci USA* 92:9643–9646.
13. Hunt, J.F., Weinkauf, S., Henry, L., Fak, J.J., McNicholas, P., Oliver, D.B., and Deisenhofer, J. (2002). Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. *Science* 297:2018–2026.
14. Xu, Z., Knafels, J.D., and Yoshino, K. (2000). Crystal structure of the bacterial protein export chaperone secB. *Nat Struct Biol* 7:1172–1177.
15. Zhou, J., and Xu, Z. (2003). Structural determinants of SecB recognition by SecA in bacterial protein translocation. *Nat Struct Biol* 10:942–947.
16. Van den Berg, B., Clemons, W.M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427:36–44.
17. Sharma, V., Arockiasamy, A., Ronning, D.R., Savva, C.G., Holzenburg, A., Braunstein, M., Jacobs, W.R., Jr., and Sacchettini, J.C. (2003). Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase. *Proc Natl Acad Sci USA* 100:2243–2248.
18. Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* 107:361–372.
19. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., III, Ban, N., and Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438:318–324.
20. Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., and Akey, C.W. (2000). The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 6:1219–1232.
21. Morgan, D.G., Menetret, J.F., Radermacher, M., Neuhof, A., Akey, I.V., Rapoport, T.A., and Akey, C.W. (2000). A comparison of the yeast and rabbit 80 S ribosome reveals the topology of the nascent chain exit tunnel, inter-subunit bridges and mammalian rRNA expansion segments. *J Mol Biol* 301:301–321.
22. de Keyzer, J., Van der Does, C., and Driessen, A.J.M. (2003). The bacterial translocase: A dynamic protein channel complex. *Cell Mol Life Sci* 60:2034–2052.
23. Veenendaal, A.K., Van der Does, C., and Driessen, A.J.M. (2004). The protein-conducting channel SecYEG. *Biochim Biophys Acta* 1694:81–95.

24. Vrontou, E., and Economou, A. (2004). Structure and function of SecA, the preprotein translocase nanomotor. *Biochim Biophys Acta* 1694:67–80.
25. Urbanus, M.L., Scotti, P.A., Froderberg, L., Saaf, A., de Gier, J.W., Brunner, J., Samuelson, J.C., Dalbey, R.E., Oudega, B., and Luirink, J. (2001). Sec-dependent membrane protein insertion: Sequential interaction of nascent FtsQ with SecY and YidC. *EMBO Rep* 2:524–529.
26. Samuelson, J.C., Jiang, F., Yi, L., Chen, M., de Gier, J.W., Kuhn, A., and Dalbey, R.E. (2001). Function of YidC for the insertion of M13 procoat protein in *Escherichia coli*: translocation of mutants that show differences in their membrane potential dependence and Sec requirement. *J Biol Chem* 276:34847–34852.
27. Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and Stuart, R.A. (2003). Yeast Oxa1 interacts with mitochondrial ribosomes: The importance of the C-terminal region of Oxa1. *EMBO J* 22:6438–6447.
28. Szyrach, G., Ott, M., Bonnefoy, N., Neupert, W., and Herrmann, J.M. (2003). Ribosome binding to the Oxa1 complex facilitates co-translational protein insertion in mitochondria. *EMBO J* 22:6448–6457.
29. Duong, F., and Wickner, W. (1997). Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. *EMBO J* 16:2756–2768.
30. Duong, F., and Wickner, W. (1997). The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. *EMBO J* 16:4871–4879.
31. Kato, Y., Nishiyama, K., and Tokuda, H. (2003). Depletion of SecDF-YajC causes a decrease in the level of SecG: Implication for their functional interaction. *FEBS Lett* 550:114–118.
32. Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., Chemama, Y., Labigne, A., *et al.* (2001). The protein-protein interaction map of *Helicobacter pylori*. *Nature* 409:211–215.
33. Nouwen, N., and Driessen, A.J.M. (2002). SecDFyajC forms a heterotetrameric complex with YidC. *Mol Microbiol* 44:1397–1405.
34. Nouwen, N., Van der Laan, M., and Driessen, A.J.M. (2001). SecDFyajC is not required for the maintenance of the proton motive force. *FEBS Lett* 508:103–106.
35. Stenberg, F., Chovanec, P., Maslen, S.L., Robinson, C.V., Ilag, L.L., von Heijne, G., and Daley, D.O. (2005). Protein complexes of the *Escherichia coli* cell envelope. *J Biol Chem* 280:34409–34419.
36. Lecker, S.H., Driessen, A.J.M., and Wickner, W. (1990). ProOmpA contains secondary and tertiary structure prior to translocation and is shielded from aggregation by association with SecB protein. *EMBO J* 9:2309–2314.
37. Gannon, P.M., Li, P., and Kumamoto, C.A. (1989). The mature portion of *Escherichia coli* maltose-binding protein (MBP) determines the dependence of MBP on SecB for export. *J Bacteriol* 171:813–818.
38. Fekkes, P., Van der Does, C., and Driessen, A.J.M. (1997). The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J* 16:6105–6113.
39. Shimizu, H., Nishiyama, K., and Tokuda, H. (1997). Expression of *gpsA* encoding biosynthetic sn-glycerol 3-phosphate dehydrogenase suppresses both the LB-phenotype of a *secB* null mutant and the cold-sensitive phenotype of a *secG* null mutant. *Mol Microbiol* 26:1013–1021.
40. Delepelaire, P., and Wandersman, C. (1998). The SecB chaperone is involved in the secretion of the Serratia marcescens HasA protein through an ABC transporter. *EMBO J* 17:936–944.

41. Baars, L., Ytterberg, A.J., Drew, D., Wagner, S., Thilo, C., Van Wijk, K.J., and de Gier, J.W. (2005). Defining the role of the *E. coli* chaperone SecB using comparative proteomics. *J Biol Chem* 281:10024–10034.
42. Knoblauch, N.T., Rudiger, S., Schonfeld, H.J., Driessen, A.J.M., Schneider-Mergener, J., and Bukau, B. (1999). Substrate specificity of the SecB chaperone. *J Biol Chem* 274:34219–34225.
43. Rajapandi, T., Dolan, K.M., and Oliver, D.B. (1991). The first gene in the *Escherichia coli secA* operon, gene X, encodes a nonessential secretory protein. *J Bacteriol* 173:7092–7097.
44. Nakatogawa, H., and Ito, K. (2001). Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Mol Cell* 7:185–192.
45. Nakatogawa, H., and Ito, K. (2002). The ribosomal exit tunnel functions as a discriminating gate. *Cell* 108:629–636.
46. Nakatogawa, H., Murakami, A., Mori, H., and Ito, K. (2005). SecM facilitates translocase function of SecA by localizing its biosynthesis. *Genes Dev* 19:436–444.
47. Chen, X., Xu, H., and Tai, P.C. (1996). A significant fraction of functional SecA is permanently embedded in the membrane. SecA cycling on and off the membrane is not essential during protein translocation. *J Biol Chem* 271:29698–29706.
48. Eichler, J., Rinard, K., and Wickner, W. (1998). Endogenous SecA catalyzes preprotein translocation at SecYEG. *J Biol Chem* 273:21675–21681.
49. Schatz, P.J., Riggs, P.D., Jacq, A., Fath, M.J., and Beckwith, J. (1989). The *secE* gene encodes an integral membrane protein required for protein export in *Escherichia coli*. *Genes Dev* 3:1035–1044.
50. Schatz, P.J., Bieker, K.L., Ottemann, K.M., Silhavy, T.J., and Beckwith, J. (1991). One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the *E. coli* secretion machinery. *EMBO J* 10:1749–1757.
51. Van der Sluis, E.O., and Driessen, A.J.M. (2006). Stepwise evolution of the Sec machinery in Proteobacteria. *Trends Microbiol* 14:105–108.
52. Braunstein, M., Brown, A.M., Kurtz, S., and Jacobs, W.R., Jr. (2001). Two nonredundant SecA homologues function in mycobacteria. *J Bacteriol* 183:6979–6990.
53. Lenz, L.L., and Portnoy, D.A. (2002). Identification of a second Listeria secA gene associated with protein secretion and the rough phenotype. *Mol Microbiol* 45: 1043–1056.
54. Bensing, B.A., and Sullam, P.M. (2002). An accessory *sec* locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. *Mol Microbiol* 44:1081–1094.
55. Braunstein, M., Espinosa, B.J., Chan, J., Belisle, J.T., and Jacobs, R. (2003). SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* 48:453–464.
56. Lenz, L.L., Mohammadi, S., Geissler, A., and Portnoy, D.A. (2003). SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proc Natl Acad Sci USA* 100:12432–12437.
57. Bensing, B.A., Gibson, B.W., and Sullam, P.M. (2004). The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. *J Bacteriol* 186:638–645.
58. Bensing, B.A., Takamatsu, D., and Sullam, P.M. (2005). Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. *Mol Microbiol* 58:1468–1481.
59. Takamatsu, D., Bensing, B.A., and Sullam, P.M. (2004). Four proteins encoded in the gspB-secY2A2 operon of *Streptococcus gordonii* mediate the intracellular glycosylation of the platelet-binding protein GspB. *J Bacteriol* 186:7100–7111.

60. Takamatsu, D., Bensing, B.A., and Sullam, P.M. (2004). Genes in the accessory sec locus of *Streptococcus gordonii* have three functionally distinct effects on the expression of the platelet-binding protein GspB. *Mol Microbiol* 52:189–203.
61. Chen, Q., Wu, H., and Fives-Taylor, P.M. (2004). Investigating the role of secA2 in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. *Mol Microbiol* 53:843–856.
62. Machata, S., Hain, T., Rohde, M., and Chakraborty, T. (2005). Simultaneous deficiency of both MurA and p60 proteins generates a rough phenotype in *Listeria monocytogenes*. *J Bacteriol* 187:8385–8394.
63. Tian, H., Boyd, D., and Beckwith, J. (2000). A mutant hunt for defects in membrane protein assembly yields mutations affecting the bacterial signal recognition particle and Sec machinery. *Proc Natl Acad Sci USA* 97:4730–4735.
64. Traxler, B., and Murphy, C. (1996). Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J Biol Chem* 271:12394–12400.
65. Van der Laan, M., Nouwen, N., and Driessen, A.J.M. (2004). SecYEG proteoliposomes catalyze the $\Delta\psi$-dependent membrane insertion of FtsQ. *J Biol Chem* 279:1659–1664.
66. Neumann-Haefelin, C., Schafer, U., Muller, M., and Koch, H.G. (2000). SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. *EMBO J* 19:6419–6426.
67. Zito, C.R., and Oliver, D. (2003). Two-stage binding of SecA to the bacterial translocon regulates ribosome-translocon interaction. *J Biol Chem* 278:40640–40646.
68. Karamyshev, A.L., and Johnson, A.E. (2005). Selective SecA association with signal sequences in ribosome-bound nascent chains: A potential role for SecA in ribosome targeting to the bacterial membrane. *J Biol Chem* 280:37930–37940.
69. Liebke, H.H. (1987). Multiple SecA protein isoforms in *Escherichia coli*. *J Bacteriol* 169:1174–1181.
70. Fekkes, P., de Wit, J.G., Van der Wolk, J.P., Kimsey, H.H., Kumamoto, C.A., and Driessen, A.J.M. (1998). Preprotein transfer to the *Escherichia coli* translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol Microbiol* 29:1179–1190.
71. Economou, A., and Wickner, W. (1994). SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* 78:835–843.
72. Schiebel, E., Driessen, A.J.M., Hartl, F.U., and Wickner, W. (1991). $\Delta\mu$ H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* 64:927–939.
73. Van der Wolk, J.P., de Wit, J.G., and Driessen, A.J.M. (1997). The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. *EMBO J* 16:7297–7304.
74. Economou, A., Pogliano, J.A., Beckwith, J., Oliver, D.B., and Wickner, W. (1995). SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* 83:1171–1181.
75. Osborne, A.R., Clemons, W.M., Jr., and Rapoport, T.A. (2004). A large conformational change of the translocation ATPase SecA. *Proc Natl Acad Sci USA* 101:10937–10942.
76. Dekker, C., de Kruijff, B., and Gros, P. (2003). Crystal structure of SecB from *Escherichia coli*. *J Struct Biol* 144:313–319.
77. Koonin, E.V., and Gorbalenya, A.E. (1992). Autogenous translation regulation by *Escherichia coli* ATPase SecA may be mediated by an intrinsic RNA helicase activity of this protein. *FEBS Lett* 298:6–8.
78. Papanikou, E., Karamanou, S., Baud, C., Frank, M., Sianidis, G., Keramisanou, D., Kalodimos, C.G., Kuhn, A., and Economou, A. (2005). Identification of the preprotein binding domain of SecA. *J Biol Chem* 280:43209–43217.

79. Kimura, E., Akita, M., Matsuyama, S., and Mizushima, S. (1991). Determination of a region in SecA that interacts with presecretory proteins in *Escherichia coli*. *J Biol Chem* 266:6600–6606.
80. Woodbury, R.L., Hardy, S.J., and Randall, L.L. (2002). Complex behavior in solution of homodimeric SecA. *Protein Sci* 11:875–882.
81. Akita, M., Shinaki, A., Matsuyama, S., and Mizushima, S. (1991). SecA, an essential component of the secretory machinery of *Escherichia coli*, exists as homodimer. *Biochem Biophys Res Commun* 174:211–216.
82. Dempsey, B.R., Economou, A., Dunn, S.D., and Shilton, B.H. (2002). The ATPase domain of SecA can form a tetramer in solution. *J Mol Biol* 315:831–843.
83. Or, E., Navon, A., and Rapoport, T. (2002). Dissociation of the dimeric SecA ATPase during protein translocation across the bacterial membrane. *EMBO J* 21:4470–4479.
84. Benach, J., Chou, Y.T., Fak, J.J., Itkin, A., Nicolae, D.D., Smith, P.C., Wittrock, G., Floyd, D.L., Golsaz, C.M., Gierasch, L.M., and Hunt, J.F. (2003). Phospholipid-induced monomerization and signal-peptide-induced oligomerization of SecA. *J Biol Chem* 278:3628–3638.
85. Musial-Siwek, M., Rusch, S.L., and Kendall, D.A. (2005). Probing the affinity of SecA for signal peptide in different environments. *Biochemistry* 44:13987–13996.
86. Bu, Z., Wang, L., and Kendall, D.A. (2003). Nucleotide binding induces changes in the oligomeric state and conformation of SecA in a lipid environment: A small-angle neutron-scattering study. *J Mol Biol* 332:23–30.
87. Bessonneau, P., Besson, V., Collinson, I., and Duong, F. (2002). The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *EMBO J* 21:995–1003.
88. Duong, F. (2003). Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. *EMBO J* 22:4375–4384.
89. Tziatzios, C., Schubert, D., Lotz, M., Gundogan, D., Betz, H., Schagger, H., Haase, W., Duong, F., and Collinson, I. (2004). The bacterial protein-translocation complex: SecYEG dimers associate with one or two SecA molecules. *J Mol Biol* 340:513–524.
90. Jilaveanu, L.B., Zito, C.R., and Oliver, D. (2005). Dimeric SecA is essential for protein translocation. *Proc Natl Acad Sci USA* 102:7511–7516.
91. de Keyzer, J., Van der Sluis, E.O., Spelbrink, R.E., Nijstad, N., de Kruijff, B., Nouwen, N., Van der Does, C., and Driessen, A.J.M. (2005). Covalently dimerized SecA is functional in protein translocation. *J Biol Chem* 280:35255–35260.
92. Karamanou, S., Sianidis, G., Gouridis, G., Pozidis, C., Papanikolau, Y., Papanikou, E., and Economou, A. (2005). *Escherichia coli* SecA truncated at its termini is functional and dimeric. *FEBS Lett* 579:1267–1271.
93. Randall, L.L., Crane, J.M., Lilly, A.A., Liu, G., Mao, C., Patel, C.N., and Hardy, S.J. (2005). Asymmetric binding between SecA and SecB two symmetric proteins: Implications for function in export. *J Mol Biol* 348:479–489.
94. Or, E., Boyd, D., Gon, S., Beckwith, J., and Rapoport, T. (2005). The bacterial ATPase SecA functions as a monomer in protein translocation. *J Biol Chem* 280:9097–9105.
95. Jilaveanu, L.B., and Oliver, D. (2006). SecA dimer cross-linked at its subunit interface is functional for protein translocation. *J Bacteriol* 188:335–338.
96. Driessen, A.J.M. (1993). SecA, the peripheral subunit of the *Escherichia coli* precursor protein translocase, is functional as a dimer. *Biochemistry* 32:13190–13197.
97. Ding, H., Hunt, J.F., Mukerji, I., and Oliver, D. (2003). *Bacillus subtilis* SecA ATPase exists as an antiparallel dimer in solution. *Biochemistry* 42:8729–8738.

98. Ding, H., Mukerji, I., and Oliver, D. (2003). Nucleotide and phospholipid-dependent control of PPXD and C-domain association for SecA ATPase. *Biochemistry* 42:13468–13475.
99. Ding, H., Mukerji, I., and Oliver, D. (2001). Lipid and signal peptide-induced conformational changes within the C-domain of *Escherichia coli* SecA protein. *Biochemistry* 40:1835–1843.
100. Chou, Y.T., Swain, J.F., and Gierasch, L.M. (2002). Functionally significant mobile regions of *Escherichia coli* SecA ATPase identified by NMR. *J Biol Chem* 277:50985–50990.
101. Ulbrandt, N.D., London, E., and Oliver, D.B. (1992). Deep penetration of a portion of *Escherichia coli* SecA protein into model membranes is promoted by anionic phospholipids and by partial unfolding. *J Biol Chem* 267:15184–15192.
102. Den Blaauwen, T., Fekkes, P., de Wit, J.G., Kuiper, W., and Driessen, A.J.M. (1996). Domain interactions of the peripheral preprotein translocase subunit SecA. *Biochemistry* 35:11994–12004.
103. Schmidt, M., Ding, H., Ramamurthy, V., Mukerji, I., and Oliver, D. (2000). Nucleotide binding activity of SecA homodimer is conformationally regulated by temperature and altered by *prlD* and *azi* mutations. *J Biol Chem* 275:15440–15448.
104. Baud, C., Karamanou, S., Sianidis, G., Vrontou, E., Politou, A.S., and Economou, A. (2002). Allosteric communication between signal peptides and the SecA protein DEAD motor ATPase domain. *J Biol Chem* 277:13724–13731.
105. Vrontou, E., Karamanou, S., Baud, C., Sianidis, G., and Economou, A. (2004). Global co-ordination of protein translocation by the SecA IRA1 switch. *J Biol Chem* 279: 22490–22497.
106. Story, R.M., Li, H., and Abelson, J.N. (2001). Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proc Natl Acad Sci USA* 98:1465–1470.
107. Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., and Wigley, D.B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97:75–84.
108. Zhou, J., and Xu, Z. (2005). The structural view of bacterial translocation-specific chaperone SecB: Implications for function. *Mol Microbiol* 58:349–357.
109. Fekkes, P., de Wit, J.G., Boorsma, A., Friesen, R.H., and Driessen, A.J.M. (1999). Zinc stabilizes the SecB binding site of SecA. *Biochemistry* 38:5111–5116.
110. Breukink, E., Nouwen, N., Van Raalte, A., Mizushima, S., Tommassen, J., and de Kruijff, B. (1995). The C terminus of SecA is involved in both lipid binding and SecB binding. *J Biol Chem* 270:7902–7907.
111. Dempsey, B.R., Wrona, M., Moulin, J.M., Gloor, G.B., Jalilehvand, F., Lajoie, G., Shaw, G.S., and Shilton, B.H. (2004). Solution NMR structure and X-ray absorption analysis of the C-terminal zinc-binding domain of the SecA ATPase. *Biochemistry* 43:9361–9371.
112. Matousek, W.M., and Alexandrescu, A.T. (2004). NMR structure of the C-terminal domain of SecA in the free state. *Biochim Biophys Acta* 1702:163–171.
113. Kimsey, H.H., Dagarag, M.D., and Kumamoto, C.A. (1995). Diverse effects of mutation on the activity of the *Escherichia coli* export chaperone SecB. *J Biol Chem* 270:22831–22835.
114. Rajapandi, T., and Oliver, D. (1994). Carboxy-terminal region of *Escherichia coli* SecA ATPase is important to promote its protein translocation activity *in vivo*. *Biochem Biophys Res Commun* 200:1477–1483.

115. Van Voorst, F., Van der Does, C., Brunner, J., Driessen, A.J.M., and de Kruijff, B. (1998). Translocase-bound SecA is largely shielded from the phospholipid acyl chains. *Biochemistry* 37:12261–12268.
116. Dapic, V., and Oliver, D. (2000). Distinct membrane binding properties of N- and C-terminal domains of *Escherichia coli* SecA ATPase. *J Biol Chem* 275:25000–25007.
117. Kinch, L.N., Saier, M.H., Jr., and Grishin, N.V. (2002). Sec61beta–a component of the archaeal protein secretory system. *Trends Biochem Sci* 27:170–171.
118. Breyton, C., Haase, W., Rapoport, T.A., Kuhlbrandt, W., and Collinson, I. (2002). Three-dimensional structure of the bacterial protein-translocation complex SecYEG. *Nature* 418:662–665.
119. Bostina, M., Mohsin, B., Kuhlbrandt, W., and Collinson, I. (2005). Atomic model of the *E. coli* membrane-bound protein translocation complex SecYEG. *J Mol Biol* 352:1035–1043.
120. Veenendaal, A.K.J., Van der Does, C., and Driessen, A.J.M. (2002). The core of the bacterial translocase harbors a tilted transmembrane segment 3 of SecE. *J Biol Chem* 277:36640–36645.
121. Beswick, V., Baleux, F., Huynh-Dinh, T., Kepes, F., Neumann, J.M., and Sanson, A. (1996). NMR conformational study of the cytoplasmic domain of the canine Sec61 gamma protein from the protein translocation pore of the endoplasmic reticulum membrane. *Biochemistry* 35:14717–14724.
122. Plath, K., Mothes, W., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* 94:795–807.
123. Wang, L., Miller, A., Rusch, S.L., and Kendall, D.A. (2004). Demonstration of a specific *Escherichia coli* SecY-signal peptide interaction. *Biochemistry* 43:13185–13192.
124. Plath, K., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (2004). Interactions between Sec complex and prepro-alpha-factor during posttranslational protein transport into the endoplasmic reticulum. *Mol Biol Cell* 15:1–10.
125. Tian, P., and Andricioaei, I. (2006). Size, motion and function of the SecY translocon revealed by molecular dynamics simulations with virtual probes. *Biophys J* 90:2718–2730.
126. Mori, H., Tsukazaki, T., Masui, R., Kuramitsu, S., Yokoyama, S., Johnson, A.E., Kimura, Y., Akiyama, Y., and Ito, K. (2003). Fluorescence resonance energy transfer analysis of protein translocase. SecYE from *Thermus thermophilus* HB8 forms a constitutive oligomer in membranes. *J Biol Chem* 278:14257–14264.
127. Van der Sluis, E., Nouwen, N., and Driessen, A.J.M. (2002). SecY-SecY and SecY-SecG contacts revealed by site-specific crosslinking. *FEBS Lett* 527:159–165.
128. Veenendaal, A.K.J., Van der Does, C., and Driessen, A.J.M. (2001). Mapping the sites of interaction between SecY and SecE by cysteine scanning mutagenesis. *J Biol Chem* 276:32559–32566.
129. Kaufmann, A., Manting, E.H., Veenendaal, A.K.J., Driessen, A.J.M., and Van der Does, C. (1999). Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. *Biochemistry* 38:9115–9125.
130. Scheuring, J., Braun, N., Nothdurft, L., Stumpf, M., Veenendaal, A.K., Kol, S., Van der Does, C., Driessen, A.J.M., and Weinkauf, S. (2005). The oligomeric distribution of SecYEG is altered by SecA and translocation ligands. *J Mol Biol* 354:258–271.
131. Nagamori, S., Nishiyama, K., and Tokuda, H. (2000). Two SecG molecules present in a single protein translocation machinery are functional even after crosslinking. *J Biochem* 128:129–137.

132. Meyer, T.H., Ménétret, J.F., Breitling, R., Miller, K.R., Akey, C.W., and Rapoport, T.A. (1999). The bacterial SecY/E translocation complex forms channel-like structures similar to those of the eukaryotic Sec61p complex. *J Mol Biol* 285:1789–1800.
133. Collinson, I., Breyton, C., Duong, F., Tziatzios, C., Schubert, D., Or, E., Rapoport, T., and Kuhlbrandt, W. (2001). Projection structure and oligomeric properties of a bacterial core protein translocase. *EMBO J* 20:2462–2471.
134. Manting, E.H., Van der Does, C., Remigy, H., Engel, A., and Driessen, A.J.M. (2000). SecYEG assembles into a tetramer to form the active protein translocation channel. *EMBO J* 19:852–861.
135. Hanein, D., Matlack, K.E., Jungnickel, B., Plath, K., Kalies, K.U., Miller, K.R., Rapoport, T.A., and Akey, C.W. (1996). Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* 87:721–732.
136. Woolf, P.J., and Linderman, J.J. (2003). Self organization of membrane proteins via dimerization. *Biophys Chem* 104:217–227.
137. Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W., von Heijne, G., Van der Does, C., Driessen, A.J.M., Oudega, B., and Luirink, J. (2000). YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J* 19:542–549.
138. Yahr, T.L., and Wickner, W.T. (2000). Evaluating the oligomeric state of SecYEG in preprotein translocase. *EMBO J* 19:4393–4401.
139. Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997). Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* 278:2123–2126.
140. Menetret, J.F., Hegde, R.S., Heinrich, S.U., Chandramouli, P., Ludtke, S.J., Rapoport, T.A., and Akey, C.W. (2005). Architecture of the ribosome-channel complex derived from native membranes. *J Mol Biol* 348:445–457.
141. Morgan, D.G., Menetret, J.F., Neuhof, A., Rapoport, T.A., and Akey, C.W. (2002). Structure of the mammalian ribosome-channel complex at 17A resolution. *J Mol Biol* 324:871–886.
142. Prinz, A., Behrens, C., Rapoport, T.A., Hartmann, E., and Kalies, K.U. (2000). Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. *EMBO J* 19:1900–1906.
143. Tani, K., Tokuda, H., and Mizushima, S. (1990). Translocation of ProOmpA possessing an intramolecular disulfide bridge into membrane vesicles of *Escherichia coli*. Effect of membrane energization. *J Biol Chem* 265:17341–17347.
144. De Keyzer, J., Van der Does, C., and Driessen, A.J.M. (2002). Kinetic analysis of the translocation of fluorescent precursor proteins into *Escherichia coli* membrane vesicles. *J Biol Chem* 277:46059–46065.
145. Raden, D., Song, W., and Gilmore, R. (2000). Role of the cytoplasmic segments of Sec61alpha in the ribosome-binding and translocation-promoting activities of the Sec61 complex. *J Cell Biol* 150:53–64.
146. Cheng, Z., Jiang, Y., Mandon, E.C., and Gilmore, R. (2005). Identification of cytoplasmic residues of Sec61p involved in ribosome binding and cotranslational translocation. *J Cell Biol* 168:67–77.
147. Nagamori, S., Nishiyama, K., and Tokuda, H. (2002). Membrane topology inversion of SecG detected by labeling with a membrane-impermeable sulfhydryl reagent that causes a close association of SecG with SecA. *J Biochem (Tokyo)* 132:629–634.
148. Nishiyama, K., Hanada, M., and Tokuda, H. (1994). Disruption of the gene encoding p12 (SecG) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature. *EMBO J* 13:3272–3277.

149. Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *EMBO J* 15:1482–1494.
150. Kalies, K.U., Rapoport, T.A., and Hartmann, E. (1998). The beta subunit of the Sec61 complex facilitates cotranslational protein transport and interacts with the signal peptidase during translocation. *J Cell Biol* 141:887–894.
151. Suzuki, H., Nishiyama, K., and Tokuda, H. (1998). Coupled structure changes of SecA and SecG revealed by the synthetic lethality of the secAcsR11 and delta secG::kan double mutant. *Mol Microbiol* 29:331–341.
152. Matsumoto, G., Mori, H., and Ito, K. (1998). Roles of SecG in ATP- and SecA-dependent protein translocation. *Proc Natl Acad Sci USA* 95:13567–13572.
153. Nishiyama, K., Suzuki, T., and Tokuda, H. (1996). Inversion of the membrane topology of SecG coupled with SecA-dependent preprotein translocation. *Cell* 85:71–81.
154. Van der Sluis, E.O., Van der Vries, E., Berrelkamp, G., Nouwen, N., and Driessen, A.J.M. (2006). Topologically fixed SecG is fully functional. *J Bacteriol* 188:1188–1190.
155. Harris, C.R., and Silhavy, T.J. (1999). Mapping an interface of SecY (PrlA) and SecE (PrlG) by using synthetic phenotypes and *in vivo* cross-linking. *J Bacteriol* 181:3438–3444.
156. Tam, P.C., Maillard, A.P., Chan, K.K., and Duong, F. (2005). Investigating the SecY plug movement at the SecYEG translocation channel. *EMBO J* 24:3380–3388.
157. Gumbart, J., and Schulten, K. (2006). Molecular dynamics studies of the archaeal translocon. *Biophys J* 90:2356–2367.
158. van der Sluis, E.O., Nouwen, N., Koch, J., de Keyzer, J., van der Does, C., Tampé, R., and Driessen, A.J.M. (2006). Identification of two interaction sites in Sec Y that are important for the functional interaction with SecA. *J Mol Biol* 361:839–849.

# 3

# *Targeting of Proteins by the Twin-Arginine Translocation System in Bacteria and Chloroplasts*

SHARON MENDEL • COLIN ROBINSON

*Department of Biological Sciences*
*University of Warwick*
*Coventry CV4 7AL*
*United Kingdom*

## I. Introduction

Translocation of proteins is an important process in essentially all living organisms. In prokaryotes, protein transport is indeed vital because numerous proteins need to be transported across the plasma membrane into the periplasm, the outer membrane (in Gram-negative organisms), or the external medium. The importance of the process can be judged by the fact that in *Escherichia coli* more than 250 proteins are transported across the plasma membrane, while 7% of the *Bacillus subtilis* proteome is believed to be exported [1]. In all cases, the cell is faced with two interesting problems: how to get proteins across the plasma membrane while keeping the compartments distinctly separate, and how to maintain membrane integrity in the process. The plasma membrane is, after all, an energy-transducing membrane that is designed to be highly impermeable even to protons, let alone proteins.

We now know that there are two main pathways for the export of proteins in most free-living bacteria: the secretory (Sec) pathway and the



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25003-6

twin-Arg translocation (Tat) pathway. They differ in fundamental respects, and particularly in the folding state of the substrate: Sec substrates are delivered into the Sec pathway in an unfolded state, and maintained in this state for the duration of the targeting pathway, while the Tat system is highly unusual in transporting its substrates in a folded state. This is an essential feature of the Tat pathway because some substrates are obliged to fold in the cytoplasm, as discussed below.

In this chapter, we focus on the Tat pathway with the aim of describing the structure, function, and mechanism of this unusual system. The Tat system is found in the cytoplasmic membranes of many bacteria, in many archaea, and in the thylakoid membranes of chloroplasts. While the overall aim is to describe the state of the art for bacterial Tat systems, the system was discovered in plant chloroplasts and much of our knowledge has emerged from studies on the thylakoid Tat system. Most Tat systems work in broadly similar ways and the chloroplast data will be described where appropriate.

## II. Basic Features of Tat Systems, Their Discovery, and Their Distribution

Early studies on the targeting of thylakoid lumen proteins in chloroplasts [2–4] showed that some proteins are transported across the thylakoid membrane by a highly unusual mechanism that is dependent on the thylakoidal $\Delta$pH but not nucleoside triphosphate (NTP) hydrolysis. These energetic requirements ruled out an involvement of the thylakoid Sec system, which is completely ATP dependent as are bacterial Sec systems. Instead, the data pointed to the presence of a completely different type of translocase. Subsequent work showed that the Sec-independent translocation system recognizes substrates with a twin-Arg motif in the signal peptide (hence the nomenclature), and the mutagenesis studies showed that the twin-Arg motif plays a critical important role. Substitution of either Arg, even by Lys, results in a complete block in translocation by the thylakoid Tat system [5, 6].

The real breakthrough in this field came with the isolation of a maize mutant, termed *hcf106* that was shown to be specifically defective in Tat-dependent protein targeting [7]. Once the *hcf106* gene was sequenced [8], it became clear that homologous genes were present in most sequenced bacterial genomes (as unassigned reading frames). This immediately indicated that a related system may operate in bacteria and this was confirmed by the isolation of *E. coli tat* mutants in 1998 [9–11]. These mutants are viable but unable to grow using anaerobic respiration pathways because several key periplasmic redox proteins are transported by this pathway (see below).

We now know that the Tat translocase is present in many but not all bacteria and is widely distributed in archaea [12, 13]. It is not present in animals or yeasts, which is perhaps surprising since mitochondria, like chloroplasts, are descended from endosymbiotic prokaryotes, but *tat* genes have been identified in some protist mitochondrial genomes. *tatC* genes are also present in plant mitochondrial genomes, although the pathway has not been characterized at all and its significance remains to be probed.

## III. *tat* Genes and Mutant Phenotypes

### A. *tat* Genes in Gram-Negative Bacteria

The initial genetic studies in *E. coli* identified four genes involved in Tat functioning: *tatA*, *tatB*, *tatC*, and *tatE* [9–11]. Consistent with a role in protein translocation, these *tat* genes were predicted to encode integral membrane proteins. Strains with disruptions in *tatB*, *tatC*, or a combination of *tatA* and *tatE* are completely defective in the translocation of a subset of periplasmic proteins, most notably those involved in anaerobic respiration pathways (see below). The *tatA*, *tatB*, and *tatC* genes form an operon with a fourth promoter-distal gene, *tatD*. The cotranscription of *tatD* with three genes known to encode essential components of the Tat export pathway initially suggested an involvement of TatD in the Tat system. However, it is now known that TatD is a soluble protein with DNase activity that has no apparent role in Tat-dependent translocation [14].

In *E. coli*, *tatA* and *tatB* encode small proteins of 9.6 and 18.4 kDa, respectively, and each contains a single transmembrane (TM) span with a cytoplasmic domain. The two proteins share significant homology but there is good evidence that they play very different roles in the overall translocation process, as explained in later sections. *tatE* encodes a TatA paralogue that is expressed at very low levels, and this subunit appears to play no specific or special role in the Tat pathway [15]. Moreover, many Gram-negative organisms contain only *tatABC* genes. *tatC* is a highly hydrophobic 28.9-kDa protein with six TM spans [16]. The essential structural features of the encoded *E. coli* Tat proteins are shown in Figure 3.1.

Detailed analyses of *tat* genes in other Gram-negative bacteria reveal a typical pattern in which the presence of *tatABC* is the norm [12]. However, many variations can be seen; as an example, *Legionella pneumophila* is a Gram-negative facultative intracellular parasite of freshwater; the *tatA* and *tatB* genes are cotranscribed, while the *tatC* gene is situated elsewhere on the chromosome.

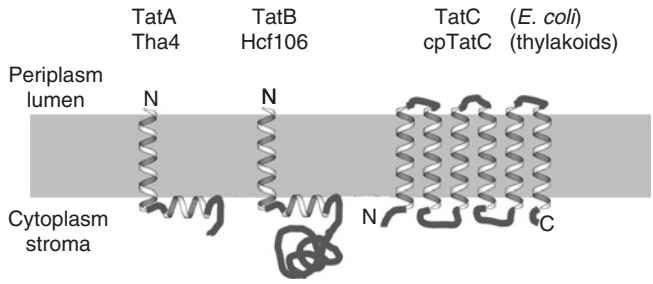



FIG. 3.1. Basic structures of Tat subunits in *E. coli* and plant thylakoids. The figure illustrates the consensus basic structures of the TatABC subunits or homologous Tha4/Hcf106/cpTatC subunits in Gram-negative bacteria and plant thylakoids, respectively. The TatA/Tha4 and TatB/Hcf106 subunits each contains three domains: A single TM span, predicted amphipathic α-helix, and soluble domain.

## B. *TAT* MUTANT PHENOTYPES AND SUBSTRATE SPECIFICITIES IN GRAM-NEGATIVE BACTERIA

Disruption of *tatA* + *tatE*, *tatB*, or *tatE* in *E. coli* leads to an inability to grow via anaerobic respiration pathways, and this reflects the primary role of the Tat system in *E. coli* and related bacteria to export periplasmic proteins bearing any of a range of redox cofactors including FeS, molybdopterin, NiFe centers, and others [9–11, 17, 18]. The essential point is that these cofactors can only be inserted into the proteins in the bacterial cytoplasm, with complex enzymatic machineries involved in some cases. This means that export of the protein in apoprotein form is not an option because it is impossible to insert any of this range of cofactors in the periplasm. This necessitates both folding and cofactor acquisition in the cytoplasm, and in turn export to the periplasm in a largely, if not fully folded form.

A detailed analysis of apparent Tat system substrates is presented in [19]. Several known Tat substrates can be assigned to specific phenotypes from studies on Tat-defective mutant strains. Most have been designated as potential Tat substrates on the basis of their signal peptides, and have yet to be formally shown to utilize the Tat pathway. Examples of known substrates in *E. coli* include trimethylamine-*N*-oxide (TMAO) reductase (TorA), formate dehydrogenase N (FDH-N), hydrogenase-2, and Cell wall amidases.

### 1. *Trimethylamine-N-Oxide Reductase, a 97-kDa Molybdoprotein Encoded by the TorA Gene [20]*

Based on sequence homologies, TorA belongs to the dimethyl sulfoxide (DMSO)/TMAO reductase family which includes the DMSO/TMAO reductases from *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*

and TMAO reductase of *Shewanella* species. These enzymes are molybdoenzymes located in the bacterial periplasm, and they interact with membrane-anchored pentahemic *c* type cytochromes, which feed electrons to the terminal enzyme [21]. It has been shown that TorD is a chaperone for TorA, encoded by the same operon in *E. coli*, which protects the TorA signal peptide from proteolysis to allow translocation of the enzyme by the Tat system [22].

### 2. *Formate Dehydrogenase N*

FDH-N of *E. coli* is a membrane-bound enzyme comprising FdnG, FdnH, and FdnI subunits organized in an $(\alpha\beta\gamma)3$ ''trimer-of-trimers'' configuration. The *fdnGHI* operon encodes proteins of 110, 32, and 20 kDa, which correspond to the subunit sizes of purified FDH-N. The FdnG subunit carries a Tat-dependent signal peptide, which directs the protein to the periplasmic side of the membrane. FDH-N can comprise up to 10% of the total membrane protein [23]. Together with nitrate reductase-A, it forms a respiratory chain transferring electrons from formate to nitrate which results in the generation of a proton motive force. FDH-N has a number of cofactors, including *bis*-molybdopterin guanine dinucleotide cofactor, selenocysteine, and a single (4Fe–4S) cluster [23, 24].

### 3. *Hydrogenase-2*

Hydrogenases catalyze the reversible oxidation of hydrogen and allow bacteria to use hydrogen as an energy source for their growth. Hydrogenases can be divided into two major superfamilies: (1) nickel-iron hydrogenases (NiFe hydrogenases) and (2) iron-only hydrogenases (Fe hydrogenases). They are generally composed of a small subunit of about 30 kDa and a large subunit of 60 kDa. All small subunits of periplasmic or membrane-bound hydrogenases contain a twin-Arg signal peptide, and the *E. coli* precursor protein has been shown to be exported by the Tat pathway [10]. The large subunits of NiFe hydrogenases show no N-terminal processing. The small and large subunits of hydrogenase-2 (HYD2) of *E. coli* are encoded by *hybO* and *hybC* of the *hybOABCDEFG* operon, respectively. HYD2 is an extrinsic membrane protein that is attached to the periplasmic side of the cytoplasmic membrane by a 5-kDa fragment of its small subunit [25].

### 4. *Cell Wall Amidases*

*E. coli tat* deletion mutants are defective in cell wall biogenesis due to the absence of the Tat substrate amidases AmiA and AmiC [26], resulting in hypersensitivity to antibiotics and detergents including SDS, as well as an

increased susceptibility to lysis by lysozyme [27]. The amidases are also responsible for the filamentous phenotype often noted in these strains pointing to a cell division defect.

### 5. *Other Phenotypes*

Recent reports have demonstrated the important contribution of the Tat pathway to the virulence of a number of organisms, including *Pseudomonas aeruginosa*, where *tat* mutations have been shown to severely affect the secretion of phospholipase C proteins, as well as motility and biofilm formation [28, 29]. Impaired secretion of Shiga toxin 1 (Stx1) and synthesis of H7 flagellin in the pathogenic *E. coli* strain O157:H7 has also been attributed to a nonfunctional Tat system [30]. Given that *tat* genes are significantly absent from human cells and that the system seems to have a central role in bacterial pathogenicity, it represents a potentially important and novel target for antimicrobial drugs. A role for Tat has also been implicated in *Agrobacterium tumefaciens*-mediated plant infections and the interaction between bacteria and host during symbiotic nitrogen fixation [31, 32].

It is generally accepted that the transport of large cofactor-containing substrates by the Tat pathway effectively represents compelling evidence for their export in a fully folded state. Strictly speaking, it has not been shown that proteins are actually transported in this manner, but other lines of evidence do support this conclusion. First, biochemical studies on thylakoids have provided direct evidence for the translocation of folded protein domains by this system [33, 34]. Second, there is good evidence for the export of hydrogenase subunits in an *oligomeric* form [25, 35]. While *E. coli* hydrogenase-2 small subunit (HybO) carries a typical twin-Arg signal peptide (as discussed above), the nickel-containing large subunit (HybC) is devoid of any known export signal. The authors concluded that the latter subunit is almost certainly exported by a ''hitchhiker'' mechanism after binding to the former Tat substrate. Finally, it has been shown that the TorA signal peptide can direct the export of green fluorescent protein (GFP) by the Tat pathway in both *E. coli* [36, 37] and *Synechocystis* sp. PCC6803 [38] and, because GFP is unable to fold into an active form when transported by the Sec pathway [39], this has been taken as further evidence for the transport of folded proteins by the Tat pathway. As an aside, this has enabled the periplasmic compartment to be visualized directly, and Figure 3.2 shows a confocal microscope image of *Synechocystis* sp. PCC6803 exporting TorA-GFP to the periplasmic compartment. The three-dimensional (3D) reconstruction shows a halo of GFP surrounding the interior of the cell, which appears red to the autofluorescence of the internal thylakoid membranes.

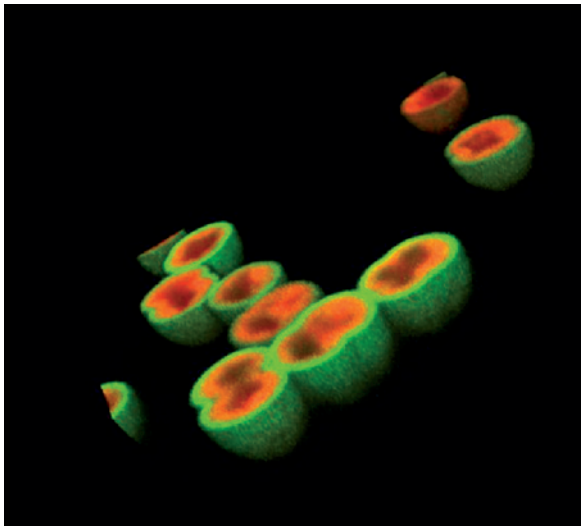


FIG. 3.2. Export of a heterologous protein, GFP, by the Tat pathway in cyanobacteria. The image shows a *Synechocystis* PCC6803 transformant expressing a construct comprising a Tat signal peptide (from *E. coli* TorA) linked to GFP. The cells were analyzed by confocal microscopy followed by 3D rendering, and the periplasmic GFP is visible as a green halo surrounding the internal thylakoids (which emit red fluorescence). Image courtesy of Anja Nenninger.

Many prokaryotes use the Tat pathway predominantly for the secretion of redox proteins, but analysis of the predicted substrates suggests that certain bacteria and archaea secrete mainly nonredox proteins via the Tat system [13, 40]. Thus, it seems likely that the Tat system is predominantly used for the export of two types of protein: those that are obliged to fold prior to export (e.g., redox proteins as discussed above) and those that cannot be transported by the Sec pathway for other reasons. Given that Sec substrates have to be transported in an unfolded form, it seems very probable that some proteins simply fold too tightly and/or rapidly for the Sec system to handle, and these would need to be transported by an alternative pathway. Figure 3.3 illustrates the basic elements of the two main protein export pathways in Gram-negative bacteria. The figure emphasizes the point that substrates are maintained in an unfolded state for the duration of the posttranslational Sec-dependent pathway. In contrast, Tat substrates are apparently able to fold in the cytoplasm, after which they can be exported either directly in a ''simple'' pathway or after binding a cofactor in the much more complex pathway undertaken by some cofactor-containing substrates.

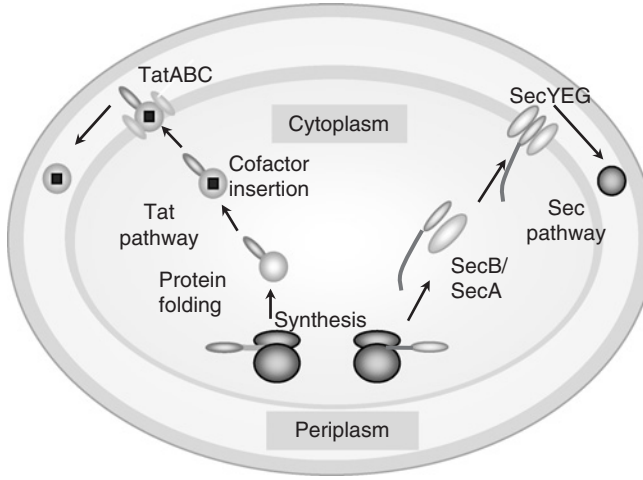



FIG. 3.3. The major pathways for protein export in *E. coli*. Proteins are exported across the plasma membrane by two main pathways. In the Sec pathway, the nascent protein is bound by chaperones such as SecB as soon as it leaves the ribosome; other chaperones interact even during translation. The protein is maintained in an unfolded conformation until it reaches the SecYEG translocon; other Sec components such as SecDF (not shown) plays ancillary roles. The driving force for translocation is provided by SecA, which hydrolyzes ATP and repeatedly pushes segments of the substrates through the SecYEG channel until translocation is complete. In the alternative, Tat-dependent route, substrate proteins are allowed to fold in the cytoplasm and in some cases this is indeed essential; key substrates include redox proteins that bind cofactors such as FeS or molybdopterin centers in the cytoplasm. It is therefore believed, although not formally proven, that these proteins are transported in a folded state by the Tat apparatus. Evidence from *in vitro* studies suggests that the proton motive force is harnessed to provide the driving force for translocation.

## C. *TAT* GENES IN PLANTS

The counterpart *tat* genes in higher plants are *tha4*, *hcf106*, and *cptatC* [8, 41]. As with the bacterial TatA and TatB subunits, Tha4 and Hcf106 share significant similarity, and assignment of a TatA/TatB-type function to the plant Tat subunits is difficult on the basis of sequence similarities alone (neither gene complements bacterial *tat* mutants; our unpublished data). However, mechanistic studies described below have clearly shown that Tha4 and Hcf106 are homologues of TatA and TatB, respectively. Mutations in any of the plant genes lead to drastic effects on thylakoid biogenesis and a seedling lethal phenotype in homozygous mutants [see 8]. This reflects a key role of the Tat system in thylakoid biogenesis; several key lumenal proteins are transported by this pathway and the chloroplast cannot achieve photosynthetic competence in the absence of a functional Tat system.

The substrate specificity of the thylakoid system is relatively well defined from a variety of studies using specific inhibitors of the Sec and Tat pathways. In addition, proteomic studies have identified numerous lumenal proteins that are strongly implicated as being Sec- or Tat dependent [42]. It appears that most lumenal proteins are transported by the Tat pathway in chloroplasts, indicating a critical role for this system in chloroplast biogenesis. Interestingly, the vast majority of chloroplast Tat substrates do not bind cofactors, suggesting that it is their folding properties that render these proteins incapable of transfer by the Sec pathway.

### D. *tat* Genes in Gram-Positive Organisms and Archaea

*tat* gene organization is very different in Gram-positive bacteria and archaea. The key difference is the absence of an apparent *tatB* gene; most Gram-positive organisms only contain obvious homologues of *tatA* and *tatC* [12, 13, 43]. The best-studied Gram-positive organism is probably *B. subtilis*, the genome of which encodes three TatA and two TatC-like proteins. Recent studies strongly indicate that they form two distinct TatAC systems. Unexpectedly, one of these systems (encoded by the *tatAd* and *tatCd* genes) appears to be used for the export of only a single substrate, namely the phosphodiesterase PhoD [44, 45]. This substrate is encoded by the same operon and all three proteins are expressed in response to phosphate starvation. The *tatAyCy* operon is believed to encode the Tat system that is used by other substrates, although only one (YwbN) has been identified to date [45]. The function of the third *tatA* gene is unclear.

Purely on the basis of these differing gene sets, it was considered likely that TatA is bifunctional in Gram-positive organisms, fulfilling both TatA- and TatB-type functions. More recently, evidence supporting this concept has emerged form studies on mutated Tat subunits in *E. coli*. In this study, Blaudeck *et al*. [46] selected for *E. coli* strains that were able to export a reporter protein using mutated, plasmid-borne *tatAC* genes (i.e., in the complete absence of TatB). Mutations in TatA were isolated that compensate for the absence of TatB and support significant levels of Tat-dependent translocation activity. All of the mutations mapped to the extreme N-terminal domain of TatA. No mutations affecting TatC were identified. These results suggest that in TatAC-type systems, the TatA protein represents a bifunctional component fulfilling both the TatA and TatB functions.

Finally, it should be emphasized that a number of organisms lack a Tat system altogether. In general, this seems to be related to a small genome size [12]. However, there are several exceptions to this, such as *Fusobacterium nucleatum* (genome size 4.33 Mb) and *Lactococcus lactis* (genome size 2.37 Mb), both of which lack a Tat system [13]. The main

source of energy of these organisms is fermentation, and, as suggested before, the presence of a Tat-system may therefore relate to a presence of extracellular redox proteins [12]. The lack of a Tat system is not necessarily related to an intracellular lifestyle. For example, the obligatorily intracellular *Rickettsia prowazekii* contains one TatA and one TatC homologue, which are predicted to export only one substrate.

## IV. The Tat Subunits: Structures and Conserved Regions

As indicated earlier, TatA and TatB are related proteins that contain a single TM span, whereas TatC contains six TM spans. Another conserved feature of the TatA and TatB subunits is the presence of a predicted amphipathic α-helix immediately after the TM helix; these basic structural features are shown in Figure 3.4. The sequences show no significant similarities to any other proteins in the database, which underlines the unusual nature of their function in this translocation pathway. However, this also means that database mining cannot provide much in the way of clues regarding the translocation mechanism. In an effort to identify

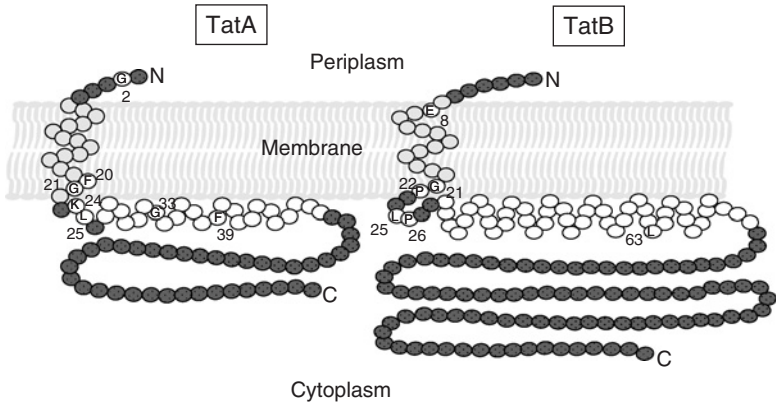



FIG. 3.4. Detailed structures of the TatA and TatB subunits. TatA and TatB are homologous proteins that nevertheless carry out very different functions in the Tat system; the majority of cellular TatB is found tightly complexes to TatC, while TatA is found both associated with the TatBC units and also as separate, relatively abundant homo-oligomeric complexes. Residues at the junction between the TM helix and predicated amphipathic helix are very highly conserved; these are F20-G21 in TatA and G21-P22 in TatB. The figure illustrates some of the residues targeted for mutagenesis; surprisingly, very few are essential for activity and only F39 appears to be critical.

essential features of the proteins, several studies have involved site-specific mutagenesis or truncation of the three Tat subunits, mainly in *E. coli*.

Truncations and point mutational analysis have showed that only the first 40 or so residues of TatA and the first 101 residues of TatB, which include the N-terminal TM span and the adjacent amphipathic helices of each protein, are necessary for the Tat function [47]. Within this region, the most conserved residues are found within the TM span and the short sequence lying between the TM and amphipathic regions. Figure 3.5 shows an alignment of these regions within TatA and TatB; very few residues are invariant even within this limited sample. TatA subunits from Gram-negative bacteria contain a highly conserved FG motif (shown in bold), while TatB sequences contain a conserved GP motif in this region (the same Gly that is conserved in TatA sequences). Confusingly, both Tha4 and Hcf106 contain FGP in this region, as do the *B. subtilis* TatA/TatB sequences; the functional significance of these motifs is thus difficult


FIG. 3.5. Conserved residues in TatA and TatB. The Figure shows an alignment of various TatA and TatB subunits from Gram-negative bacteria, with the TM span and predicted amphipathic regions indicated. Identical residues are shown by asterisks and conserved residues by dots. Conserved FG and GP motifs are hallmarks of TatA/TatB, respectively, from Gram-negative bacteria, but note the presence of both motifs (i.e., FGP) in the TatAd subunit from *B. subtilis* and thylakoid Hcf106 subunit.

to judge. One difference between the subunit types is the presence of an almost invariant Glu at the periplasm/membrane border within TatB (residue 8 within the *E. coli* sequence). Its presence within the TM span region of TatB subunits has promoted interest because of the potential for this residue to engage in proton transfer pathways. However, translocation is inhibited but certainly not blocked by substitution of this residue [48, 49], raising questions about its significance. Further mutagenesis of TatA and TatB has confirmed that the predicted amphipathic region plays a critical role and this does appear to be a central feature within each subunit [48–50]. The only residue shown to be indispensable for translocation, F39, is located in this region.

The TatC subunit has also been subjected to a fair amount of site-specific mutagenesis and the results have again been rather disappointing [51, 52]. There are relatively few conserved residues in this subunit too and only a handful are essential of those probed so far. The data have so far revealed little about the actual translocation mechanism.

## V. Structures of Tat Complexes

Several efforts have been made to study the organization of Tat complexes, and while some conflicting results have been published, there now seems to be agreement on the overall nature of these complexes, at least in *E. coli*. Overexpression of the *E. coli tatABC* operon, with a Strep II tag attached to the C-terminus of TatC, led to the purification of a TatABC complex in the nonionic detergent digitonin [53]. This complex eluted as a ca. 650-kDa assembly containing roughly equimolar amounts of TatA, TatB, and TatC. Given that the total mass of *E. coli* TatABC is 57 kDa, results point to the presence of multiple copies of each subunit. The large digitonin micelle presumably accounts for a considerable proportion of the size estimates determined by gel filtration chromatography. It has been shown that TatB and TatC are present in 1:1 stoichiometry in the *E. coli* TatABC complex, and that a translational fusion of these subunits is fully active. The TatB and C subunits therefore appear to act as a structural and functional unit within this complex.

More recently, similar TatABC complexes from *Salmonella typhimurium* and *A. tumefaciens* have been purified after expression of the respective *tatABC* operons in *E. coli* [54]. Large complexes of about 650 kDa containing only TatABC were purified from detergent-solubilized *E. coli* membranes, confirming that these proteins participate in a complex and again raising the possibility that these three proteins carry out the central functions in the translocation process. The combined molecular mass of TatB

and TatC is 42.7 kDa, so that five to seven TatBC units would amount to 240–330 kDa; the presence of one TatA subunit (9.8 kDa) per TatBC would increase the size of the complexes to between 290 and 400 kDa. Single particle electron microscope data revealed an asymmetric complex with dimensions of up to 10 × 13 nm, containing stain-excluding domains that may correspond to TatBC units (or TatABC units).

While *S. typhimurium* was chosen as an example of an organism that is closely related to *E. coli*, the *A. tumefaciens* Tat subunit sequences are far less similar to those of the *E. coli* subunits and the Tat system transports a very different range of subunits. The TatABC complex has a similar overall organization to those of the *E. coli* and *S. typhimurium* complexes, although it is slightly smaller. These results suggest that a TatABC complex of ca. 400–500 kDa may be a consensus feature within Gram-negative organisms. A different study involved single particle analysis of complexes that contained only TatA and B, in the form of double-ring structures [55]. However, the absence of TatC raises questions about the significance of this complex because TatC has been shown to bind firmly to TatB in several other studies.

In the thylakoid system, the Tat complexes have not been purified and the available data come from blue native gel studies where the complexes are identified using immunoblotting. It has been shown [56] that Hcf106 and cpTatC (homologues of TatB and TatC, respectively) are tightly linked within a large complex (ca. 700 kDa in blue native gels). This is consistent with the coupling of these subunits in bacterial TatABC complexes, but one apparent difference is the absence of the TatA homologue, Tha4, in this complex.

Although TatA has been shown to be firmly linked to TatBC in purified bacterial TatABC complexes, the vast majority of TatA does not copurify with TatBC and instead forms separate homo-oligomeric complexes within the plasma membrane. Studies have shown that these complexes have unusual features that may be highly relevant to the translocation mechanism. In one study [57], blue native gel electrophoresis showed that these complexes are remarkably heterogeneous, ranging in size from about 50 kDa to well over 500 kDa. The same study showed that the TatABC complex is discrete in nature, running as a tight band of about 370 kDa—consistent with the gel filtration studies since the detergent micelle is effectively replaced by Coomassie dye in the blue native system, usually resulting in a smaller, more accurate size estimation.

A higher-resolution 3D image of a Tat complex was recently obtained using random canonical tilt electron microscopy. A complex consisting almost entirely of TatA forms pore-like structures of various sizes that appear gated or blocked on the cytoplasmic side of the inner membrane [58].

The structural data provide support for the idea that the TatA protein forms the channel (or at least part of the channel) through which substrates are translocated across the membrane.

## VI. Tat Signal Peptides

Substrates for the Tat pathway bear N-terminal signal peptides that comprise three distinct domains (Figure 3.6): a polar N-terminal (N-) domain, hydrophobic core (H-) domain, and C-terminal domain terminating in an Ala-Xaa-Ala consensus motif specifying cleavage by signal peptidase. A twin-Arg motif is almost invariably present at the junction of the N- and C-domains, and it is this motif that gave rise to the Tat nomenclature.

Perhaps surprisingly, Tat signal peptides are very similar in some respects to Sec-type signal peptides, which have the same basic three-domain organization [59]. Despite this overall conservation, signal peptides found in Tat substrates do have several distinguishing features when compared to those found in Sec substrates. While Sec-type signals have no real primary


FIG. 3.6. Tat signal peptides. The figure shows representative Tat signals from the Gram-negative bacterium, *E. coli*, the cyanobacterium *Synechocystis* PCC6803, and the archaeon *Halobacterium* sp. NRC-1. Twin-Arg motifs and other charged residues are shown in bold, and the hydrophobic domains underlined. Note that the signals for most of the *E. coli* proteins have been confirmed as Tat dependent; the archaeal and cyanobacterial signals are merely predicted to specify Tat-dependent targeting (although the close resemblances strongly suggest this function, see text).

sequence conservation apart from the consensus Ala-Xaa-Ala motif at the C-terminus, detailed studies of bacterial Tat signals have revealed a consensus SRRxFLK motif (where x is a polar amino acid) around the twin-Arg motif [17]. The twin-Arg motif is found in chloroplast Tat signal peptides but Phe and Lys are not enriched at the +2/+4 positions. Mutagenesis studies have shown that the twin-Arg motif is absolutely critical in chloroplast Tat signals; substitution of either Arg, even by Lys, results in a complete block in translocation [5]. In *E. coli*, however, mutation of the almost invariant twin-Arg to twin-Lys completely abolishes Tat-specific export but the conservative substitution of a single Arg usually affects only the rate of translocation [60–63]. Given these results, it is perhaps unsurprising that one natural Tat substrate has been shown to contain only a single Arg (the TtrB subunit of *Salmonella enterica* tetrathionate reductase) [64]. Apparently, the chloroplast Tat system is unusually selective with its absolute requirement for twin-Arg in its substrates' signal peptides.

The significance of the commonly found +2 Phe and +4 Lys has also been probed by mutagenesis [61]. Substitution of the +2 Phe in the SufI signal peptide was found to inhibit the rate of translocation, suggesting an important role and explaining the conservation of this residue among bacterial Tat signals. Presumably, other determinants ensure efficient translocation in those signals lacking the +2 Phe. However, substitution of the +4 Lys actually enhanced the rate of translocation, suggesting a possible braking role that may tie in with ''proofreading'' functions in the overall translocation process. Prokaryotic Tat signals from *E. coli*, the archaeon *Halobacterium*, and the cyanobacterium *Synechocystis* PCC6803 are shown in Figure 3.5. *Synechocystis* signal peptides are from known periplasmic proteins [65] but none have been confirmed as authentic Tat signals; neither have the *Halobacterium* signals been experimentally analyzed in any detail.

To date, most studies have analyzed Tat signal peptides from Gram-negative bacteria. However, *Streptomyces coelicolor*, a Gram-positive eubacterium, contains the most abundant predicted Tat-dependent signal sequences among the 84 bacterial genomes analyzed [13]. Using *Streptomyces lividans*, Li *et al.* [66] have shown the functionality of nine putative Tat-dependent signal peptides that were initially predicted using the Tatscan program [67]. The results showed that the net charge in the N-domain of this Tat-dependent signal peptide had little effect on Xylanse C (XlnC) secretion. Whatever mutations were introduced in the signal sequence, they did not abolish XlnC secretion. Moreover, mutations in the signal peptidase recognition site had no effect on the precursor processing rate, suggesting that the signal peptidase might be different from the one involved in the maturation of the Sec-dependent precursor where mutations in the signal peptidase recognition site abolished precursor processing in *S. lividans* [68].

## VII. The Tat Mechanism

### A. The Translocation Mechanism Used by the *E.coli* and Thylakoid Tat Systems

It is widely accepted, though perhaps not formally proven, that the Tat system transports large proteins in a folded form, but the actual translocation mechanism is still poorly understood. Nevertheless, important advances have been made in recent years, and several salient points appear to have been resolved.

It is first appropriate to mention early work on the thylakoid system, largely involving the use of *in vitro* assays in which radiolabeled precursor protein was generated in a cell-free system and incubated with isolated thylakoids. These studies [2–4] showed that the system is dependent on the thylakoid ΔpH but not NTP hydrolysis; unusual requirements because virtually all mainstream protein transporters rely on either ATP or GTP hydrolysis at some stage. Much more recently, efficient *in vitro* assays have been developed using inverted *E. coli* membrane vesicles [69], and there again appears to be a requirement for the proton motive force. It is thus possible that this is the driving force for translocation by the Tat system. An elegant study by Alder and Theg [70] showed that the system may be costly to operate; careful measurements of the prevailing ΔpH revealed a ''cost'' of about 30,000 protons per protein transported. Assuming that folded proteins were being transported through relatively large pores, it seems reasonable to assume that this figure may represent the actual cost of protein translocation (the proton:protein stoichiometry) plus any proton leakage during the translocation process.

More recent studies have used cross-linking techniques to specifically study both the initial substrate–translocase interaction and the subsequent translocation process. Using isolated thylakoids, it was shown that substrates bind to the Hcf106 and cpTatC subunits under energy-depleted conditions that prevented further translocation [56]. This strongly suggested a role in substrate binding for the Hcf106-cpTatC complex, and subsequent work using *E. coli*-inverted membrane vesicles has corroborated this work; specific cross-links to TatB and TatC were observed [69]. It therefore appears clear that the TatBC heterodimer contains the substrate-binding site. The subsequent stages of the translocation process are ill defined but another cross-linking study made the potentially important observation that Tha4 (TatA homologue) was only found cross-linked to the Hcf106-cpTatC complex in the presence of both substrate and a ΔpH across the membrane [71]. This finding provides evidence that the separate Tha4 complex is only

recruited to the substrate-binding complex after binding of substrate and in the presence of a proton motive force.

There is scant information on other elements of the reaction and as yet we do not know how the Tat ''supercomplex'' assembles, how it forms a (presumably) large translocation pore, or how proteins are actually transported through. There may also be additional facets of the translocation mechanism that need detailed study if we are to understand this system. For example, there are intriguing indicators of a ''proofreading'' mechanism, whereby the Tat system only transports proteins when they have folded into the correct tertiary structure. Such a system makes sense, of course; it would be disastrous if cofactor-containing proteins were exported before the cofactor was inserted. Some proteins appear to circumvent this problem by the use of system-specific chaperones that bind to the signal peptide and prevent transport until a correct conformation is reached and the chaperone is somehow displaced. The best-studied example is TMAO reductase (TorA) which is encoded by the *torCAD* operon. TorC is a partner protein (a membrane-bound quinol dehydrogenase), and Jack *et al.* [72] have shown that TorD is a cytoplasmic chaperone that performs two roles. The first role is to assist cofactor insertion into the TorA apoprotein, and the second role appears to involve the specific binding of TorD to the TorA signal peptide. There is strong evidence that this system prevents the export of the apoprotein until cofactor insertion is complete.

However, the Tat system may have a more general, inbuilt ability to preferentially recognize folded proteins. *In vivo* studies [73, 74] demonstrate an ability of the Tat pathway to selectively choose between at least some properly folded and misfolded proteins *in vivo*, and these studies suggest the existence of a ''folding quality control'' mechanism.

Figure 3.7 shows a proposed Tat mechanism in *E. coli*. The figure depicts translocation of a simple substrate (with no cofactors) comprising a passenger protein with a cleavable Tat signal peptide that forms a partially helical structure in apolar environments. The diagram reflects current thinking on a two-stage mechanism in which the binding of substrate to TatBC triggers the recruitment of the separate TatA complex to form the full translocation machine. However, it is important to note that the later stages of the translocation process are very poorly understood.

## B. A Different Translocation Mechanism in *B. subtilis*?

The first identified Tat substrate in the Gram-positive *B. subtilis* was PhoD, a Sec phosphodiesterase, and it has been shown [44, 75, 76] that the *tatAd* and *tatCd* genes, colocalized with *phoD* in an operon, were essential

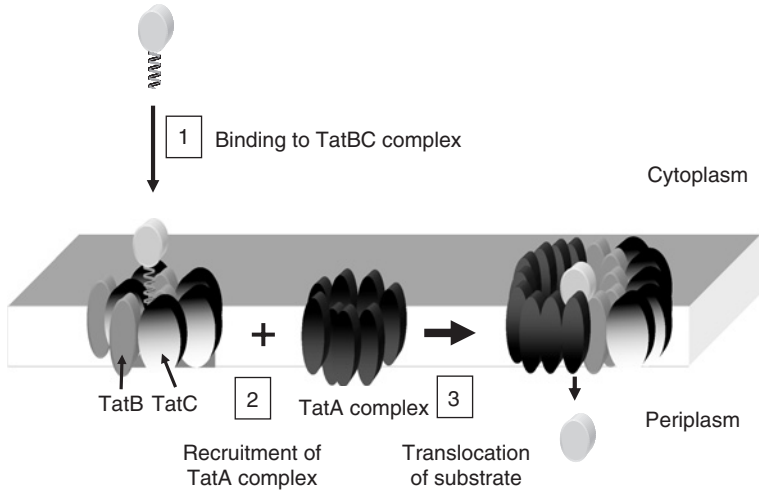



FIG. 3.7. The Tat mechanism. Substrates for the Tat system are translocated posttranslationally, at least in the characterized cases. The diagram depicts translocation of a substrate comprising a passenger protein with a cleavable Tat signal peptide that forms a partially helical structure in apolar environments. TatB and TatC form a heterodimeric unit that serves as the primary binding site for substrate (step 1: binding). Note that the large size of the TatBC complex (~370 kDa or more) suggests the presence of several binding sites. Some TatA is bound to this TatBC core, but this is not depicted in the diagram, and most is in the form of a separate homo-oligomeric complex. Binding of the substrate to the TatBC subunits is proposed to trigger the recruitment of this TatA complex (step 2: assembly), resulting in the formation of the active translocon and translocation of the substrate (step 3). The mechanism at this point is particularly ill defined; the translocation channel may be formed from multiples of TatA or may result from fusion of the two complexes as shown in this figure. Other mechanisms are also conceivable.

and sufficient to export PhoD, while the second copy of *tatC* (*tatCy*) was not required for PhoD export. TatCy, together with TatAy, forms a second TatAC translocase mediating the export of YwbN [45]. Two minimal Tat translocases are thus active in *B. subtilis*, each composed of specific TatA and TatC molecules.

More recent studies on *B. subtilis* have raised interesting issues in this field. It has been reported [77] that soluble TatAd is able to bind the twin-Arg signal peptide of prePhoD prior to membrane integration, effectively acting as a guidance factor to mediate targeting to the membrane. Evidence has been presented to suggest that TatCd serves as a receptor for TatAd–prePhoD complex, stabilizing TatAd in the membrane probably by assisting the formation of the protein-conducting channel to mediate prePhoD transport [77].

## REFERENCES


1. Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S., and Van Dijl, J.M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* 64:515–547.
2. Mould, R.M., and Robinson, C. (1991). A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. *J Biol Chem* 266:12189–12193.
3. Cline, K., Ettinger, W.F., and Theg, S.M. (1992). Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J Biol Chem* 267:2688–2696.
4. Klosgen, R.B., Brock, I.W., Herrmann, R.G., and Robinson, C. (1992). Proton gradient-driven import of the 16 kDa oxygen-evolving complex protein as the full precursor protein by isolated thylakoids. *Plant Mol Biol* 18:1031–1034.
5. Chaddock, A.M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R.G., Klosgen, R.B., and Robinson, C. (1995). A new type of signal peptide: Central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. *EMBO J* 14:15–22.
6. Henry, R., Carrigan, M., McCaffrey, M., Ma, X., and Cline, K. (1997). Targeting determinants and proposed evolutionary basis for the Sec and the delta pH protein transport systems in chloroplast thylakoid membranes. *J Cell Biol* 136:823–832.
7. Voelker, R., and Barkan, A. (1995). Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J* 14:3905–3914.
8. Settles, A.M., Yonetani, A., Baron, A., Bush, D.R., Cline, K., and Martienssen, R. (1997). Sec-independent protein translocation by the maize Hcf106 protein. *Science* 278: 1467–1470.
9. Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A., and Turner, R.J. (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93:93–101.
10. Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C., and Palmer, T. (1998). Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J* 17:3640–3650.
11. Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C., and Palmer, T. (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* 273:18003–18006.
12. Yen, M.R., Tseng, Y.H., Nguyen, E.H., Wu, L.F., and Saier, M.H.J. (2002). Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* 177:441–450.
13. Dilks, K., Rose, R.W., Hartmann, E., and Pohlschroder, M. (2003). Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* 185:1478–1483.
14. Wexler, M., Sargent, F., Jack, R., Stanley, N.R., Bogsch, E.G., Robinson, C., Berks, B.C., and Palmer, T. (2000). TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in Sec-independent protein export. *J Biol Chem* 275:16717–16722.
15. Sargent, F., Stanley, N.R., Berks, B.C., and Palmer, T. (1999). Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* 274:36073–36082.
16. Behrendt, J., Standar, K., Lindenstrauss, U., and Bruser, T. (2004). Topological studies on the twin-arginine translocase component TatC. *FEMS Microbiol Lett* 234:303–308.

17. Berks, B.C. (1996). A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* 22:393–404.
18. Santini, C.L., Ize, B., Chanal, A., Muller, M., Giordano, G., and Wu, L.F. (1998). A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* 17:101–102.
19. Berks, B.C., Palmer, T., and Sargent, F. (2003). The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* 47:187–254.
20. Silvestro, A., Pommier, J., Pascal, M.C., and Giordano, G. (1989). The inducible trimethylamine N-oxide reductase of *Escherichia coli* K12: Its localization and inducers. *Biochim Biophys Acta* 999:208–216.
21. Pommier, J., Méjean, V., Giordano, G., and Iobbi-Nivol, C. (1998). TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in *Escherichia coli*. *J Biol Chem* 273:16615–16620.
22. Genest, O., Seduk, F., Ilbert, M., Mejean, V., and Iobbi-Nivol, C. (2006). Signal peptide protection by specific chaperone. *Biochem Biophys Res Commun* 339:991–995.
23. Enoch, H.G., and Lester, R.L. (1975). The purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. *J Biol Chem* 250:6693–6705.
24. Punginelli, C., Ize, B., Stanley, N.R., Stewart, V., Sawers, G., Berks, B, and Palmer, T. (2004). mRNA secondary structure modulates translation of Tat-dependent formate dehydrogenase N. *J Bacteriol* 186:6311–6315.
25. Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L.F. (1999). Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem* 274:13223–13228.
26. Ize, B., Stanley, N.R., Buchanan, G., and Palmer, T. (2003). Role of the *Escherichia coli* Tat pathway in outer membrane integrity. *Mol Microbiol* 48:1183–1193.
27. Stanley, N.R., Findlay, K., Berks, B.C., and Palmer, T. (2001). *Escherichia coli* strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope. *J Bacteriol* 183:139–144.
28. Ochsner, U.A., Snyder, A., Vasil, A.I., and Vasil, M.L. (2002). Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc Natl Acad Sci USA* 99:8312–8317.
29. Voulhoux, R., Ball, G., Ize, B., Vasil, M.L., Lazdunski, A., Wu, L.F., and Filloux, A. (2001). Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* 20:6735–6741.
30. Pradel, N., Ye, C., Livrelli, V., Xu, J., Joly, B., and Wu, L.F. (2003). Contribution of the twin-arginine translocation system to the virulence of enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* 71:4908–4916.
31. Ding, Z., and Christie, P.J. (2003). *Agrobacterium tumefaciens* twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol* 185:760–771.
32. Meloni, S., Rey, L., Sidler, S., Imperial, J., Ruiz-Argüeso, T., and Palacios, J.M. (2003). The twin-arginine translocation (Tat) system is essential for *Rhizobium-legume* symbiosis. *Mol Microbiol* 48(5):1195–1207.
33. Hynds, P.J., Robinson, D., and Robinson, C. (1998). The Sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J Biol Chem* 273:34868–34874.
34. Clark, S.A., and Theg, S.M. (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol Biol Cell* 8:923–934.
35. Chanal, A., Santini, C.L., and Wu, L.F. (2003). Specific inhibition of the translocation of a subset of *Escherichia coli* TAT substrates by the TorA signal peptide. *J Mol Biol* 327:563–570.

36. Thomas, J.D., Daniel, R.A., Errington, J., and Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in *Escherichia coli*. *Mol Microbiol* 39:47–53.
37. Santini, C.L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C., and Wu, L.F. (2001). Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem* 276:8159–8164.
38. Spence, E., Sarcina, M., Ray, N., Moller, S.G., Mullineaux, C.W., and Robinson, C. (2003). Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium *Synechocystis* PCC6803. *Mol Microbiol* 48:1481–1489.
39. Feilmeier, B.J., Iseminger, G., Ischroeder, D., Webber, H., and Phillips, G.J. (2000). Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *J Bacteriol* 182:4068–4076.
40. Bolhuis, A. (2002). Protein transport in the halophilic archaeon Halobacterium sp. NRC-1: A major role for the twin-arginine translocation pathway? *Microbiology* 148:3335–3346.
41. Walker, M.B., Roy, L.M., Coleman, E., Voelker, R., and Barkan, A. (1999). The maize tha4 gene functions in Sec-independent protein transport in chloroplasts and is related to hcf106, tatA, and tatB. *J Cell Biol* 147:267–276.
42. Peltier, J.B., Emanuelsson, O., Kalume, D.E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D.A., Soderberg, L., Roepstorff, P., von Heijne, G., and van Wijk, K.J. (2002). Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. *Plant Cell* 14:211–236.
43. Wu, L.F., Ize, B., Chanal, A., Quentin, Y., and Fichant, G. (2000). Bacterial twin-arginine signal peptide-dependent protein translocation pathway: evolution and mechanism. *J Mol Microbiol Biotechnol* 2:179–189.
44. Pop, O., Martin, U., Abel, C., and Muller, J.P. (2002). The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of *Bacillus subtilis* form an autonomous Tat translocation system. *J Biol Chem* 277:3268–3273.
45. Jongbloed, J.D., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S., and van Dijl, J.M. (2004). Two minimal Tat translocases in *Bacillus*. *Mol Microbiol* 54:1319–1325.
46. Blaudeck, N., Kreutzenbeck, P., Müller, M., Sprenger, G.A., and Freudl, R. (2005). Isolation and characterization of bifunctional *Escherichia coli* TatA mutant proteins that allow efficient Tat-dependent protein translocation in the absence of TatB. *J Biol Chem* 280:3426–3432.
47. Lee, P.A., Buchanan, G., Stanley, N.R., Berks, B.C., and Palmer, T. (2002). Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation. *J Bacteriol* 184:5871–5879.
48. Barrett, C.M.L., Mathers, J.E., and Robinson, C. (2003). Identification of key residues within the *Escherichia coli* TatAB subunits. *FEBS Lett* 537:42–46.
49. Hicks, M.G., de Leeuw, E., Porcelli, I., Buchanan, G., Berks, B.C., and Palmer, T. (2003). The *Escherichia coli* twin-arginine translocase: conserved residues of TatA and TatB family components involved in protein transport. *FEBS Lett* 539:61–67.
50. Barrett, C.M.L., and Robinson, C. (2005). Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin-arginine translocase. *FEBS J* 272:2261–2275.
51. Allen, S.C.H., Barrett, C.M.L., Ray, N., and Robinson, C. (2002). Essential cytoplasmic domains in the *Escherichia coli* TatC protein. *J Biol Chem* 277:10362–10366.
52. Buchanan, G., De Leeuw, E., Stanley, N.R., Wexler, M., Berks, B.C., Sargent, F., and Palmer, T. (2002). Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol Microbiol* 43:1457–1470.

53. Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C.M., and Robinson, C. (2001). TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *J Biol Chem* 276:20213–20219.
54. Oates, J., Mathers, J., Mangels, D., Kuhlbrandt, W., Robinson, C., and Model, K. (2003). Consensus structural features of purified bacterial TatABC complexes. *J Mol Biol* 330:277–286.
55. Sargent, F., Gohlke, U., *et al.* (2001). Purified components of the *Escherichia coli* Tat protein transport system form a double-layered ring structure. *Eur J Biochem* 268:3361–3367.
56. Cline, K., and Mori, H. (2001). Thylakoid Delta pH-dependent precursor bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J Cell Biol* 154:719–729.
57. Oates, J., Barrett, C.M.L., Barnett, J.P., Byrne, K.G., Bolhuis, A., and Robinson, C. (2005). The *Escherichia coli* twin-arginine translocation apparatus incorporates a distinct form of TatABC complex, spectrum of modular TatA complexes and minor TatAB complex. *J Mol Biol* 346:295–305.
58. Gohlke, U., Pullan, L., McDevitt, C.A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H.R., and Berks, B.C. (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci USA* 102:10482–10486.
59. Driessen, A.J., Fekkes, P., and Van der Wolk, J.P. (1998). The Sec system review. *Curr Opin Microbiol* 1:216–222.
60. Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R., and Sprenger, G.A. (1999). The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of *Zymomonas mobilis* depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Eur J Biochem* 263:543–551.
61. Stanley, N.R., Palmer, T., and Berks, B.C. (2000). The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* 275:11591–11596.
62. Buchanan, G., Sargent, F., Berks, B.C., and Palmer, T. (2001). Genetic screen for suppressors of *Escherichia coli* Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. *Arch Microbiol* 177(1):107–112.
63. DeLisa, M.P., Samuelson, P., Palmer, T., and Georgiou, G. (2002). Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* 277:29825–29831.
64. Hinsley, A.P., Stanley, N.R., Palmer, T., and Berks, B.C. (2001). A naturally occurring bacterial Tat signal peptide lacking one of the ''invariant'' arginine residues of the consensus targeting motif. *FEBS Lett* 497:45–49.
65. Fulda, S., Huang, F., Nilsson, F., Hagemann, M., and Norling, B. (2000). Proteomics of *Synechocystis* sp. *strain* PCC6803. Identification of periplasmic proteins in cells grown at low and high salt concentrations. *Eur J Biochem* 267:5900–5907.
66. Li, H., Faury, D., and Morosoli, R. (2006). Impact of amino acid changes in the signal peptide on the secretion of the Tat dependent xylanase C from *Streptomyces lividans*. *FEMS Microbiol Lett* 255:268–274.
67. Li, H., Jacaues, P., *et al.* (2005). Determining the functionality of putative Tat dependent signal peptides in *Streptomyces coelicolor* A3(2) by using two different reporter proteins. *Microbiology* 151:2189–2198.
68. Morosoli, R., and Dupont, C. (1999). Secretion of xylanase A2 in *Streptomyces lividans*: dependence on signal peptides length, number and composition. *FEMS Microbiol Lett* 179:437–445.
69. Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H.G., Brunner, J., and Muller, M. (2003). Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* 12:937–946.

70. Alder, N.N., and Theg, S.M. (2003). Energetics of protein transport across biological membranes. A study of the thylakoid Delta pH-dependent/cp Tat pathway. *Cell* 112:231–242.
71. Mori, H., and Cline, K. (2002). A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase. *J Cell Biol* 157:205–210.
72. Jack, R.L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004). Coordinating assembly and export of complex bacterial proteins. *EMBO J* 23:3962–3972.
73. DeLisa, M.P., Tullman, D., and Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci USA* 100:6115–6120.
74. Fisher, A.C., Kim, W., and DeLisa, M.P. (2006). Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein Sci* 15:449–458.
75. Jongbloed, J.D., Antelmann, H., Hecker, M., Nijland, R., Bron, S., Airaksinen, U., Pries, F., Quax, W.J., van Dijl, J.M., and Braun, P.G. (2002). Selective contribution of the twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*. *J Biol Chem* 277:44068–44078.
76. Jongbloed, J.D., Martin, U., Antelmann, H., Hecker, M., Tjalsma, H., Venema, G., Bron, S., van Dijl, J.M., and Muller, J. (2000). TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *J Biol Chem* 275:41350–41357.
77. Schreiber, S., Stenge, R., Westermann, M., Volkmer-Engert, R., Pop, O., and Jörg, P.M. (2006). Affinity of TatCd for TatAd elucidates its receptor function in the *Bacillus subtilis* twin arginine translocation (Tat) translocase system. *J Biol Chem* 281:19977–19984.

This page intentionally left blank

# 4

# *YidC: A Protein with Multiple Functions in Bacterial Membrane Biogenesis*

NIL CELEBI • ROSS E. DALBEY

*Department of Chemistry*
*The Ohio State University*
*Columbus*
*OH 43210, USA*

## I. Introduction

The assembly of proteins into biological membranes is a vital process essential for all cell life. This process requires newly synthesized proteins to be efficiently targeted to the membrane, inserted, and folded into the membrane protein's three-dimensional structure. The insertion and assembly of membrane proteins is important because membrane proteins perform so many critical physiological functions in the cell, including metabolite exchange, signal transduction, communication, and energy transduction. In this capacity, the membrane proteins function in diverse roles such as membrane channels, cell surface receptors, ATPases, sugar transporters, and proton pumps. In addition, integral membrane proteins are the most popular drug targets of pharmaceutical companies.

Whether in eukaryotes or prokaryotes, each membrane protein typically begins its existence in the cytoplasm of the cell. From the cytoplasm, proteins are inserted into the membrane by the action of membrane-embedded protein translocases. In eukaryotes, the Sec61$\alpha\beta\gamma$ translocase mediates the insertion of proteins into the endoplasmic reticulum (ER)



DOI: 10.1016/S1874-6047(07)25004-8

membrane (for review see [1]). Targeting to the membrane occurs cotranslationally and requires the signal recognition particle (SRP) and the SRP receptor (for review see [2]).

The major pathway for membrane insertion into the bacterial membrane is the Sec pathway (Figure 4.1), which is related to the Sec-type pathway in the ER system. The Sec translocase is composed of the membrane integral components SecYEG, and SecDFyajC complexes, in addition to the peripheral membrane component SecA (for review see [3]). SecYEG forms the protein-conducting channel in which secreted and membrane proteins are exported across or integrated into the cell membrane [4, 5]. SecY is homologous to the Sec61$\alpha$ protein and SecE is homologous to the Sec61$\gamma$ protein; SecG and Sec61$\beta$ are functional analogues but do not show sequence homology. SecA functions as the motor ATPase (SecA not shown in Fig. 4.1; for review see [6]), promoting translocation of hydrophilic segments in defined steps during translocation across the channel. Many of the Sec-dependent proteins are believed to be targeted to the membrane by the SRP pathway (for review see [7]). SRP is composed of fifty-four homologue (Ffh) and a 4.5S RNA in *Escherichia coli*. The SRP-bound membrane proteins are delivered to the membrane by the interaction of SRP with its SRP receptor. The details of this process are still under investigation.

Operating in parallel to the Sec pathway is the YidC pathway that is used to direct a subset of membrane proteins into the bacterial membrane (Figure 4.1). Originally, these YidC-dependent inner membrane proteins were believed to insert into the membrane spontaneously. In bacteria,

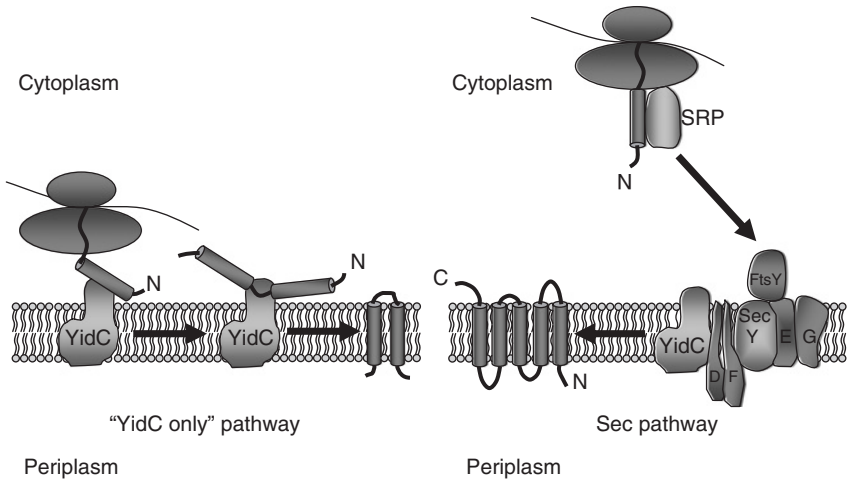



FIG. 4.1. Insertion of bacterial membrane proteins by the Sec and YidC pathways.

the YidC insertase can work on its own or with the Sec machinery [8] to insert proteins into the membrane. YidC family members are also found in eukaryotes where they are involved in membrane protein biogenesis [9–11]. In this chapter, we will discuss the role of YidC in membrane protein biogenesis, emphasizing the current knowledge of the bacterial system.

The initial discovery of the YidC/Oxa1/Alb3 family of proteins comes from studies in mitochondria on the assembly of energy transducing membrane protein components. The mitochondrial YidC homologue Oxa1p is critical for the assembly of the cytochrome *c* oxidase complex [9, 11] and for the assembly of the $F_1F_0$ ATP synthase [12, 13]. Specifically, Oxa1p was shown to play an important insertion role in the membrane biogenesis of subunit II of cytochrome *c* oxidase [14, 15]. Other proteins that require Oxa1p for efficient insertion into the mitochondrial inner membrane are Oxa1p itself, Cox1, Cox3, cytochrome *b*, and ATP6 [16].

The Oxa1-dependent pathway has been coined the ''conservative pathway'' because some of the insertion requirements for membrane proteins inserting from the matrix into the inner membrane of mitochondria are analogous to those in bacterial membrane insertion from the cytoplasm into the inner membrane [17, 18]. These Oxa1-dependent proteins include nuclearly encoded proteins that are imported from the cytoplasm into the mitochondrial matrix, then inserted into the inner membrane. In addition, certain mitochondrially encoded proteins follow this conservative pathway and are inserted directly into the inner membrane from the matrix by an Oxa1-dependent mechanism. In this latter cotranslational pathway, Oxa1p functions as a ribosome receptor, in addition to a translocase [19, 20].

The chloroplast also contains YidC homologues called the Alb3 proteins, which are similar to those found in bacteria and mitochondria. Alb3 is essential for thylakoid biogenesis. An *alb3* null mutant in *Arabidopsis thaliana* shows a severe defect on chloroplast development, resulting in an Albino phenotype [21]. Alb3 is required for insertion into the thylakoid membrane of a subset of light-harvesting chlorophyll-binding proteins (LHCB), which play important roles in transferring excitation energy to the reaction centers [22, 23]. Another chloroplast Oxa1 homologue called Alb4.1 was discovered to be necessary to maintain the correct ultra structure of the chloroplasts. This suggests that it is involved in chloroplast biogenesis [24].

In *Chlamydomonas reinhardtii*, two Alb3 proteins are involved in chloroplast thylakoid membrane biogenesis [25]. Alb3.2 is essential for the survival of *Chlamydomonas* where it plays a critical role in the thylakoid membrane assembly of the photosystem I and photosystem II reaction centers that play critical roles in light harvesting and primary electron transfer in photosynthesis. On the other hand, Alb3.1 is not essential for

the cell [26] but is required for the efficient assembly of photosystem II and for maintaining thylakoid membrane levels of LHCB [27].

## II. The YidC Pathway

The most extensively studied proteins that are inserted by the Sec-independent YidC pathway in bacteria are the M13 phage procoat protein and Pf3 coat protein (for review see [28]). Procoat is synthesized with a leader sequence and contains one transmembrane segment in the 50-residue mature domain of the protein. The membrane insertion of procoat was unaffected in conditionally lethal SecY or SecA strains [29]. Insertion of procoat could occur under drastically low SecYEG levels when SecE was depleted in the cell [30]. Similarly, the single-spanning Pf3 phage coat protein was proposed to be inserted by a Sec-independent mechanism. Its insertion was unaffected when the function of SecY or SecA was impaired using temperature-sensitive mutant strains [31]. Both Pf3 coat and procoat are inserted by an SRP-independent pathway.

Examples of endogenous *E. coli* proteins that were proposed to insert Sec independently are KdpD sensor [32], melibiose permease [33], the KcsA $K^+$ channel [34], and subunit c of $F_1F_0$ ATP synthase [35–38]. In each of these cases, membrane insertion occurred under conditions where the SecE component was severely depleted. At a minimum, the results suggested that the insertion pathway of these proteins is fundamentally different from the prototypic Sec translocase mechanism.

In recent years, the discovery of the Oxa1 conservative pathway in mitochondria and Oxa1 homologues in bacteria greatly impacted the ideas in the bacterial field. Before this discovery, the mitochondrial membrane insertion and the bacterial membrane insertion fields were generally independent of each other. This was for the most part because mitochondria do not have the Sec or SRP system. The discovery of Oxa1 started the flow of information between the mitochondrial and bacterial fields. This led to the discovery of YidC, which plays a critical role in the membrane insertion of Sec-independent proteins [39, 40]. Before this time, it was assumed that the Sec-independent membrane proteins were inserted by a ''spontaneous'' insertion pathway [41–44].

The first evidence that YidC plays a role in the insertion of proteins into the bacterial membrane came from studies showing interaction of YidC with the membrane protein FtsQ during its insertion into the membrane [8]. In addition, YidC is copurified with the Sec translocase.

The proof of YidC's role in insertion *in vivo* was the demonstration that the membrane insertion of the M13 procoat protein [39, 45] and the Pf3

coat protein [46] was strongly inhibited when YidC was depleted. This led to a paradigm shift in scientists' views on Sec-independent membrane protein insertion. Proteins, such as the M13 procoat protein and Pf3 coat, which were thought to insert into membranes spontaneously, are now known to be inserted by the action of the YidC insertase.

Generally, YidC is not required for protein export, although it is required for the export of certain lipoproteins [47]. In *in vivo* studies, the role of YidC was evaluated using the YidC depletion strain JS7131, where the chromosomal copy of *yidC* is inactivated and another copy of the gene is introduced under the control of the *araBAD* promoter. YidC is depleted by growth of the bacteria in glucose for various lengths of time.

Studies indicate that the *E. coli* YidC primarily plays a specific role in energy transduction processes similar to that of Oxa1 in mitochondria [48]. YidC depletion causes a marked induction in the so-called phage shock protein PspA [48] which is induced when the membranes are damaged and the proton motive force cannot be maintained [49]. Depletion of YidC also inhibits the cellular ATPase activity of the $F_1F_0$ ATP synthase and the activity of cytochrome bo oxidase. Indeed, the membrane levels of subunit c of the $F_1F_0$ ATP synthase and CyoA (subunit II) of cytochrome bo oxidase were reduced as a function of YidC depletion [48].

The key question is whether the YidC insertase operates by itself to promote the insertion of Sec-independent substrates. Currently, it seems likely that it does, but one cannot rule out that other proteins may regulate insertion or make membrane assembly more efficient for Sec-independent proteins. *In vitro*, YidC is sufficient for membrane insertion of Sec-independent proteins; purified denatured Pf3 coat protein can insert into YidC proteosomes [50]. The amount of Pf3 coat inserted into the YidC proteoliposomes depends on the amount of YidC in the lipid vesicle and is time dependent. YidC, alone in the vesicles, is also sufficient to insert subunit c of the $F_1F_0$ ATP synthase [36]. Subunit c does not insert into SecYEG vesicles nor does the addition of SecYEG to the vesicles containing YidC stimulate insertion.

## III. Sec–YidC Pathway

In addition to functioning alone, YidC also works with the SecYEG/SecDF translocase where it facilitates insertion and lateral integration of membrane proteins. Subunit a of the $F_1F_0$ ATP synthase [37] and some Sec-dependent procoat proteins require YidC for efficient insertion [45]. Some membrane proteins, such as CyoA, interact with YidC first and then with the Sec translocase during their insertion. The N-terminal periplasmic domain of CyoA is

inserted by the YidC pathway and the large C-terminal domain of CyoA is inserted by the Sec pathway [51–53]. CyoA seems to insert into the membrane in an N-terminal to C-terminal direction [51, 52]. Generally, YidC promotes the insertion of small hydrophobic proteins but only marginally facilitates the translocation of large C-terminal domains of membrane proteins such as leader peptidase [39], FtsQ [54], or CyoA [51].

Sec-dependent membrane proteins have been shown to contact the integral YidC protein, prior to integration into the lipid bilayer [8, 39, 54, 55]. The FtsQ hydrophobic segment interacts sequentially with SecY and then YidC [54]. Beck *et al*. [56] showed that YidC interacts with the hydrophobic regions of the polytopic membrane protein MtlA in a nonsequential manner. Taken together, these studies show that YidC must be positioned close to the SecYEG channel such that it can interact with the hydrophobic regions during the membrane topogenesis process.

Where might YidC be located in relation to the SecYEG complex? A potential answer to this question comes from the high resolution X-ray structure of the Sec complex [4]. The recent structure of the Secαβγ (SecYEG) complex shows that the Sec complex has an hourglass structure and is composed of 10 transmembrane helices from Sec61α-subunit (SecY) and 1 from each Sec61γ(SecE) and Sec61β (SecG) (Figure 4.2A) [4]. A channel with a constriction point that is occluded by a helix runs through the middle of the monomeric Sec61abg (SecYEG). This channel

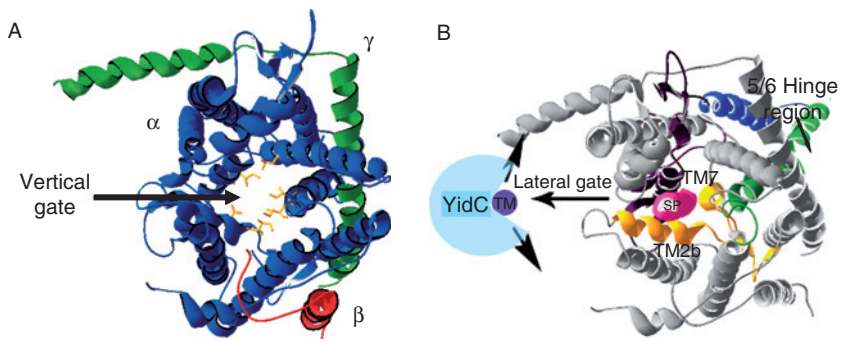



FIG. 4.2. Structure of the *M. janashii* Sec61αβγ. (A) Protein fold of Sec61αβγ showing the pore residues (view from the cytosol). Sec61α is shown in blue, Sec61γ in green, Sec61β in red, the six hydrophobic residues lining the pore ring are shown in yellow. Note that the plug is removed. (B) Putative lateral gate (TM2b/TM7) interface region (view from the cytosol). Sec61α TM5 and TM6 are indicated in green and blue, respectively. The hinge region between these transmembrane segments is indicated. TM2 is indicated in yellow and TM7 in purple. The hydrophobic core of the signal peptide (SP) is indicated in pink. SP is intercalated between TM2b and part of TM7. *Modified from* Nature *[4] with permission.*

running through the center of the SecYEG protomer is the vertical gate (Figure 4.2A; plug is omitted from structure). In addition, the structure shows that SecYEG has an intramembrane gate region called the lateral gate that could open allowing hydrophobic regions of membrane proteins to exit the channel (Figure 4.2B). Thus, it would follow that YidC would be close to the lateral gate region.

A very provocative model has been proposed based on cryo-EM studies of the *E. coli* SecYEG protein-conducting channel bound to a translating ribosome [5, 57]. In this model, a nascent membrane protein is believed to be inserted into the membrane by the action of two SecYEG translocases. One of the SecYEG protomers plays an active role in translocating hydrophilic domains across the membrane whereas the other protomer functions in the integration of the hydrophobic segments into the lipid bilayer. This model explains why the oligomeric state of the translocase in *E. coli* is a dimer [58, 59].

In certain cases, YidC may possess a discrete activity in removing the transmembrane regions of membrane protein substrates from the SecYEG channel. Evidence for this comes from experiments where the overproduction of Sec-dependent leader peptidase [45] or TatC [35] blocks the export of the precursor of OmpA (pro-OmpA) and the peptidoglycan-associated lipoprotein (PAL) only under YidC depletion conditions. The block in the export of these proteins is due to the jamming of the Sec channel by the overexpressed protein when YidC is depleted. Consequently, this data suggests that YidC is required for efficient removal of Sec-dependent proteins from the Sec translocase channel.

Another proposed function of YidC is that it constitutes an assembly site for transmembrane regions of polytopic membrane proteins during membrane biogenesis, possibly being influential in the proper folding of these membrane proteins [56]. Indeed, it has been reported that YidC may be involved in the folding of the Sec-dependent lactose permease (LacY) [60]. The lactose transporter can be inserted into the membrane in the absence of YidC. However, when YidC is depleted, the folding of LacY is perturbed. The misfolding of LacY was deduced using two monoclonal antibodies that recognize folded regions in the polar loops. This observation is intriguing and novel because it suggests that YidC is a membrane-embedded chaperone involved in the folding of polytopic membrane proteins.

## IV. YidC Substrates

Currently, the known substrates that require YidC for membrane biogenesis are a diverse group of proteins (Figure 4.3). These include M13 procoat and Pf3 coat proteins, subunits a and c of the $F_1F_0$ ATP synthase,

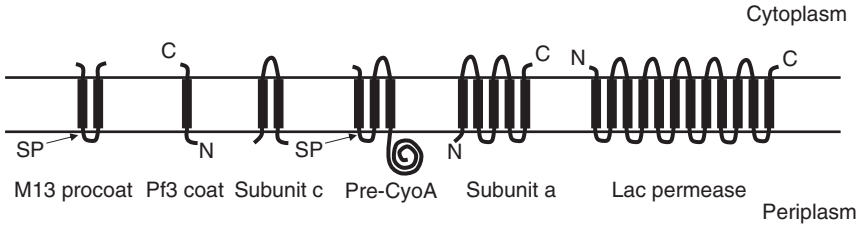



FIG. 4.3. Substrates of YidC. Procoat, Pf3 coat, and subunit c of the $F_1F_0$ ATP synthase are Sec-independent substrates. Subunit a of the $F_1F_0$ ATP synthase and Lac Permease are Sec-dependent YidC-dependent substrates. The N-terminal domain of Pre-CyoA inserts by the Sec-independent pathway, while the large C-terminal domain inserts by the Sec-dependent pathway. The black rectangles represent membrane spanning regions; SP depicts the signal peptidase cleavage site.

CyoA subunit of cytochrome *bo* oxidase, and Lac permease. The substrate LacY requires YidC for the correct folding but not for insertion [60].

## V. YidC Family of Proteins

Thus far, the best evidence for evolutionary conservation of YidC–Oxa1–Alb3 pathway in bacteria, chloroplasts, and mitochondria comes from expression studies in heterologous systems where YidC homologues from different systems can substitute for each other. The chloroplast Alb3 protein in *Arabidopsis* can substitute for YidC in *E. coli* [61]. The conserved 5 transmembrane domain of Alb3 was fused to the first 57 amino acids of YidC, and this chimeric protein was shown to complement the growth defect of the YidC-depletion strain. This protein functions in the insertion of both Sec-independent and Sec-dependent proteins in *E. coli*. Similarly, the mitochondrial Oxa1 can function in bacteria [62]. The Oxa1 protein was able to catalyze the membrane insertion of the Sec-independent procoat protein and subunit c of the $F_1F_0$ ATP synthase, but could not substitute for YidC as a chaperone in the folding of the LacY. From this study, it can be concluded that the Sec-independent function of YidC is evolutionarily conserved and essential. Conversely, the *E. coli* YidC protein was able to substitute for Oxa1 in mitochondria when the ribosome-binding C-terminal matrix domain of Oxa1 was appended to YidC [63].

Examination of the amino acid sequences of the YidC proteins reveals that there are conserved sequences localized to the hydrophobic regions of the proteins. Before we describe the conserved motifs in this family, we will review the membrane topologies of these proteins. In Gram-negative bacteria, most YidC family members are predicted to have six

transmembrane segments with the N- and C-termini facing the cytoplasm (Figure 4.4). The *E. coli* YidC spans the membrane six times with the first transmembrane segment functioning as a signal sequence to translocate the large N-terminal domain of YidC to the periplasmic space [64]. In Gram-positive bacteria, YidC is predicted to span the membrane five times (Figure 4.4). Most of these YidC homologues in Gram-positive bacteria are lipoproteins and are synthesized with a signal peptide and are processed by lipoprotein signal peptidase [65]. The topology of these YidC homologues is similar to that predicted for the mitochondrial Oxa1 and the chloroplast Alb3 protein. Generally, Gram-positive bacteria have shorter N-terminal extracytoplasmic domains than Gram-negative bacteria.

A YidC sequence alignment for five Gram-negative bacteria, five Gram-positive bacteria, *A. thaliana* Alb3, and *S. cerevisiae* Oxa1 is shown in Figure 4.5. Amino acids corresponding to the predicted transmembrane segments of *E. coli* YidC are bold and underlined [64]. The regions of consensus sequences observed in the entire YidC/Oxa1/Alb3 family [66] are shaded. These conserved sequences are found primarily in three regions within the protein (Figure 4.5) [66]. In the *E. coli* YidC, the first conserved region is within transmembrane segment 2 and has the consensus sequences GHy(AS)$(LIV)_4$T$(LIV)_3$(KR)$(LIV)_4$P(LIV). The second consensus sequence is localized to transmembrane segment 3: $KX_3$HyNP$X_2$GCLP$(LIV)_3$Q(LIV)P$(LIV)_3$AHy(YF). The third region has the conserved residues $(Hy)_2$P$(Hy)_6$XHyXHyP(SA)G$Hy_3$YW$Hy_2$(SGN)N$Hy_2$(ST) within transmembrane segments 5 and 6. Here ''Hy'' and ''X'' represent hydrophobic residue and any residue, respectively. Alternative possibilities for a

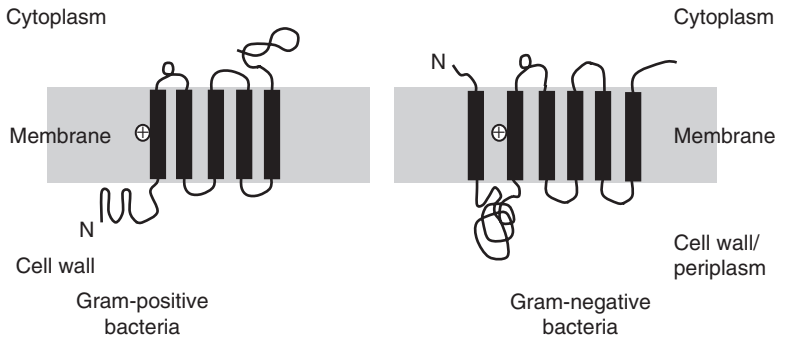



FIG. 4.4. Predicted topology of YidC family members. Proteins of the YidC family are predicted to span the membrane six or five times. The *E. coli* YidC has been shown by alkaline phosphatase fusion methods to span the membrane six times [64]. The *Bacillus subtilis YidC* homologue SpoIIIJ and YqjG are predicted to span the membrane five times [65]. SpoIIIJ and YqjG are made in a precursor form with a cleavable signal peptide.

```
E.coli                -MDSQRNLLVIALLFVSFMIWQAWEQDK-NPQPQAQQTTQTTTT---------------- 42
S.typhimurium         -MDSQRNLLVIALLFVSFMIWQAWEQDK-NPQPQTQQTTQTTTT---------------- 42
P.fluorescens         -MDIKRTILIVALAIVSYVMVLKWNQDYGQAALPTQNVASSTTTSGLPDTATGNNAAASD 59
H.influenzae          -MDSRRSLLVLALIFISFLVYQQWQLDK-NPPVQTEQTTSITATS--------------- 43
H.pylori              MDKNNNNLRLILAIALSFLFIALYSYFFQKPNKTTTQTTKQETTN--------------- 45
B.subtilis.Oxa1       ------------------------------------------------------------
S.pneumoniae          ------------------------------------------------------------
L.lactis              ------------------------------------------------------------
S.pyogenes            ------------------------------------------------------------
S.mutans.YidC2        ------------------------------------------------------------
A.thaliana.Alb3       ------------------------------------------------------------
S.cerevisiae.Oxa1     ------------------------------------------------------------

E.coli                ----AAGSAADQGVPASGQ-GKLISVKTDVLDLTINTRGGDVEQALLPAYPKELN-STQP 96
S.typhimurium         ----AAGSAADQGVPASGQ-GKMITVKTDVLDLTINTRGGDVEQALLPAYPKELG-SNEP 96
P.fluorescens         DIPRAASDTSAPAETPVAASKDLIQIKTDVLDLSIDPQGGDVAQLTLPLYPRRQDRPDVP 119
H.influenzae          -DVPASSPSNSQAIADSQTRGRIITLENDVFRLKIDTLGGDVISSELLKYDAELD-SKTP 101
H.pylori              --------NHTATSPNAPNAQHFSTTQTTPQENLLSTISFEHARIEIDSLGRIKQVYLKD 97
B.subtilis.Oxa1       ------------------------------------------------------------
S.pneumoniae          ------------------------------------------------------------
L.lactis              ------------------------------------------------------------
S.pyogenes            ------------------------------------------------------------
S.mutans.YidC2        ------------------------------------------------------------
A.thaliana.Alb3       ------------------------------------------------------------
S.cerevisiae.Oxa1     ------------------------------------------------------------

E.coli                FQLLETSPQFIYQAQSGLTGRDGPDNPANGPRPLYNVEKDAYVLAEGQNELQVPMTYTDA 156
S.typhimurium         FQLLETTPQFIYQAQSGLTGRDGPDNPANGPRPLYNVEKEAFVLADGQNELQVPMTYTDA 156
P.fluorescens         FQLFDNGGERTYLAQSGLIGTNGPDANPAG-RPIYSSEKKTYQLADGQDKLVVDLKFS-K 177
H.influenzae          FELLKDTKEHIYIAQSGLIGKNGIDTRSG--RAQYQIEGDNFKLAEGQESLSVPLLFE-K 158
H.pylori              KKYLTPKQKGFLEHVGHLFSSKENAQPPLKELPLLAADKLKPLEVRFLDPTLNNKAFNTP 157
B.subtilis.Oxa1       ------------------------------------------------------------
S.pneumoniae          ------------------------------------------------------------
L.lactis              ------------------------------------------------------------
S.pyogenes            ------------------------------------------------------------
S.mutans.YidC2        ------------------------------------------------------------
A.thaliana.Alb3       ------------------------------------------------------------
S.cerevisiae.Oxa1     ------------------------------------------------------------

E.coli                AGNTFTKTFVLKRGDYAVNVNYNVQNAGEKPLEISSFGQLKQSITLPPHLDTGSSNFALH 216
S.typhimurium         AGNTFTKTFVFKRGDYAVNVNYSVQNAGEKPLEVSTFGQLKQSVNLPPHRDTGSSNFALH 216
P.fluorescens         DGVNYIKRFTLKRGLYDVTVTYLIDNQSAQPWSGSMFAQLKRDASADP---SSTTATGTA 234
H.influenzae          DGVTYQKIFVLKRGSYDLGVDYKIDNQSGQAIEVEPYGQLKHSIVES------SGNVAMP 212
H.pylori              YSASKTTLGPNEQLVLTQDLGTLSIIKTLTFYDDLHYDLKIAFKSPNNLIPSYVITNGYR 217
B.subtilis.Oxa1       ------------------------------------------------------------
S.pneumoniae          ------------------------------------------------------------
L.lactis              ------------------------------------------------------------
S.pyogenes            ------------------------------------------------------------
S.mutans.YidC2        ------------------------------------------------------------
A.thaliana.Alb3       ------------------------------------------------------------
S.cerevisiae.Oxa1     ------------------------------------------------------------

E.coli                TFRGAAYSTPDEKYEKYKFDTIADNENLNISSKGGWVAMLQQYFATAWIPHNDGTNNFYT 276
S.typhimurium         TFRGAAYSTPDEKYEKYKFDTIADNENLNVSSKGGWVAMLQQYFATAWIPRNDGTNNFYT 276
P.fluorescens         TYLGAALWTSSEPYKKVSMKDMDKA-QLKETVTGGWVAWLQHYFVTAWVAPKGENNIVQT 293
H.influenzae          TYTGGAYSSSETNYKKYSFSDMQDN-NLSIDTKAGWVAVLQHYFVSAWIPNQDVNNQLYT 271
H.pylori              PVADLDSYTFSGVLLENSDKKIEKIEDKDAKEIKRFSNTLFLSSVDRYFTTLLFTKDPQG 277
B.subtilis.Oxa1       ------------------------------------------------------------
S.pneumoniae          ------------------------------------------------------------
L.lactis              ------------------------------------------------------------
S.pyogenes            ------------------------------------------------------------
S.mutans.YidC2        ------------------------------------------------------------
A.thaliana.Alb3       -MARVLVSSPSSFFGSPLIKPSSSLRHSGVGGGGTAQFLPYRSNNNKLFTTSTTVRFSLN 59
S.cerevisiae.Oxa1     ------------MFKLTSRLVTSRFAASSRLATARTIVLPRPHPSWISFQAKRFNSTGPN 48
```

FIG. 4.5. (*Continued*)

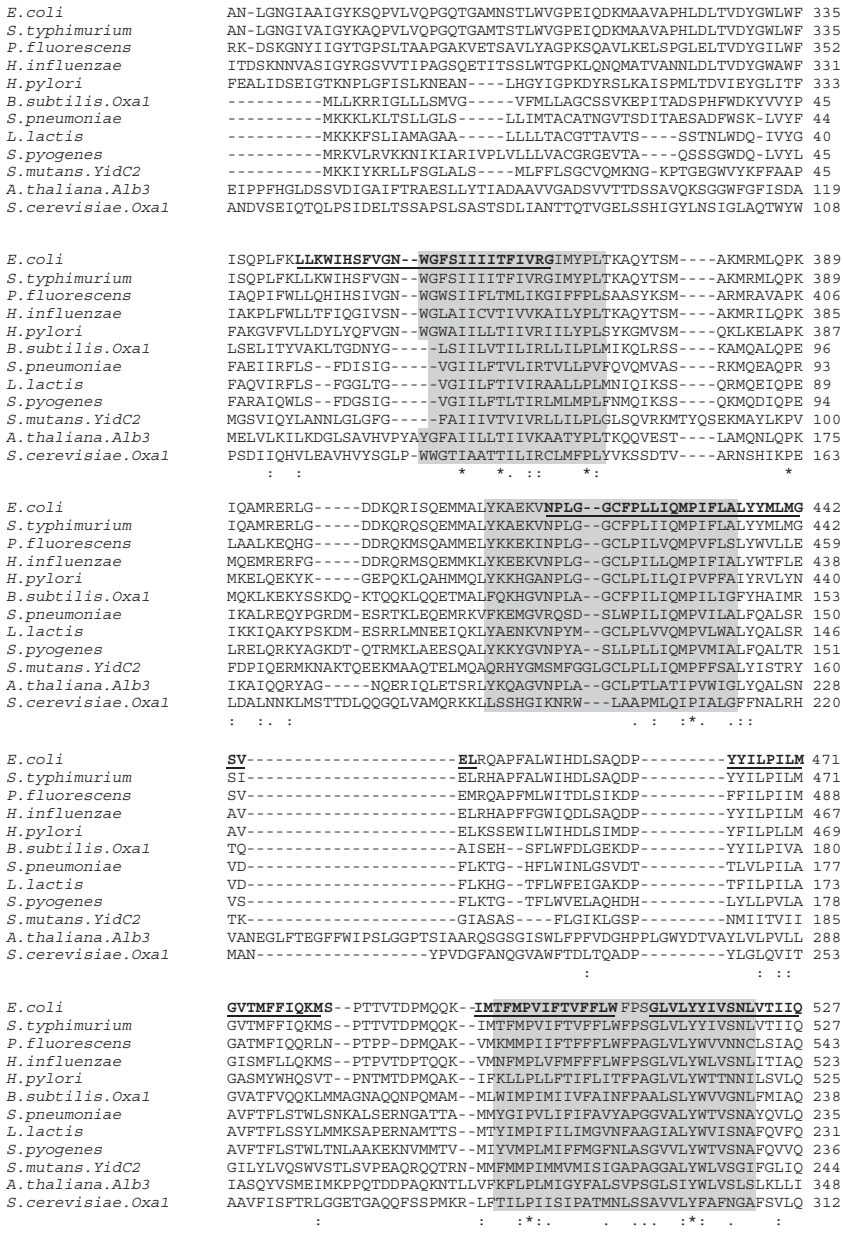



FIG. 4.5. (*Continued*)

```
E.coli                QQLIYRG----------------------------------LEKRGLHSREKKKS--- 548
S.typhimurium         QQLIYRG----------------------------------LEKRGLHSREKKKS--- 548
P.fluorescens         QWYITRK----------------------------------IEAATKKAEA------- 560
H.influenzae          QQLIYRG----------------------------------LEKKGLHSRKK------ 541
H.pylori              QLIINKVLE--------------------------------NKKRMHAQNKKEH--- 547
B.subtilis.Oxa1       TFLIKGP----------------------------DIKKNPEPQKAGGKKK------ 261
S.pneumoniae          TYFLNNPFKI-------------------IAEREAVVQAQKDLENRKRKAKKKAQKTK- 274
L.lactis              TMLLANPYKI-------------------IAAREAKVQVEKDKIKAREKALKKARKK-- 269
S.pyogenes            LLLLNNPFKI-------------------IAERQRLANEEKERRLRERRARKKAMKRK- 275
S.mutans.YidC2        QLITNHIIKP-------------------KLRKQIDEEFKKNPPKPFKSNARKDITPQA 284
A.thaliana.Alb3       ------------------------------------------------------------
S.cerevisiae.Oxa1     TMILRNKWVRSKLKITEVAKPRTPIAGASPTENMGIFQSLKHNIQKARDQAERRQLMQDN 372

E.coli                ----------------------------
S.typhimurium         ----------------------------
P.fluorescens         ----------------------------
H.influenzae          ----------------------------
H.pylori              ----------------------------
B.subtilis.Oxa1       ----------------------------
S.pneumoniae          ----------------------------
L.lactis              ----------------------------
S.pyogenes            ----------------------------
S.mutans.YidC2        NNDKKLITSKKQKSNRNAGKQRHHKQ---- 310
A.thaliana.Alb3       ----------------------------
S.cerevisiae.Oxa1     EKKLQESFKEKRQNSKIKIVHKSNFINNKK 402
```

FIG. 4.5. Amino acid sequences of bacterial YidC family members. Shown are the sequences for *E. coli*, *Escherichia coli*; *S. typhimurium*, *Salmonella typhimurium*; *P. fluorescens*, *Pseudomonas fluorescens*; *H. influenzae*, *Haemophilus influenzae*; *H. pylori*, *Helicobacter pylori*; *B. subtilis*, *Bacillus subtilis*; *S. pneumoniae*, *Streptococcus pneumoniae*; *L. lactis*, *Lactococcus lactis*; *S. pyogenes*, *Streptococcus pyogenes*; *S. mutans*, *Streptococcus mutans*; *A. thaliana*, *Arabidopsis*, *thaliana*; *S. cerevisiae*, *Saccharomyces cerevisiae*. "*" means that the residues or nucleotides in that column are identical in all sequences in the alignment. ":" means that conserved substitutions have been observed, according to the nature of the amino acid (hydrophobic, hydrophilic, acidic or basic). "." means that semi-conserved substitutions are observed.

single residue are shown in parentheses. Subscripts indicate the number of consecutive residues. These regions containing the consensus sequences are very likely to be important for the structure and function of the YidC family proteins since they are conserved throughout evolution.

Structure–function studies on the *E. coli* YidC were performed to examine the regions that are important for function [67]. These studies show that 90% of the large N-terminal periplasmic domain can be deleted without any effect on function. Likewise, the C-terminal cytoplasmic domain can be deleted without any effect on YidC function. On the other hand, the conserved transmembrane regions 2, 3, and 6 are important for activity [67]. Transmembrane segments 4 and 5 do not appear to be as important since they can be replaced by transmembrane sequences from leader peptidase. The data showing the importance of hydrophobic regions of YidC is consistent with its function as a scaffold for the insertion, binding, and folding of the hydrophobic regions of membrane proteins.

So far, there is limited structural information on the YidC membrane insertase. The oligomeric structure of YidC within intact membranes is still unclear although a portion of the purified YidC appears as a dimer on a blue native polyacrylamide gel [68], whereas Oxa1 in mitochondria is a tetramer [69]. The way in which transmembrane segments are packed to form the membrane-embedded structure and the location of the substrate-binding region of YidC are not known.

## VI. Concluding Remarks and Outlook

In summary, YidC1/Oxa1/Alb3 family members play an important role in the membrane biogenesis of proteins. YidC performs dual functions in the process of membrane protein insertion. It can promote the insertion of a subset of proteins that do not require the Sec translocase, as well as functioning in conjunction with the core Sec translocase to assist in the insertion of proteins into the membrane. For Sec-dependent proteins, YidC possibly helps fold membrane proteins during the insertion/assembly process.

Exactly how YidC functions in the insertion of proteins is not yet known. It may be a membrane chaperone assisting in the insertion and folding of proteins as they assemble into the membrane. Such a membrane chaperone may be involved in the folding of transmembrane helices to promote the helix–helix interactions within polytopic membrane proteins. Factors that determine whether a substrate goes by YidC or the Sec pathway are not clear. In order to understand how YidC works at a molecular level in membrane protein biogenesis, it is absolutely essential to solve its structure.

### References

1. Stephenson, K. (2005). Sec-dependent protein translocation across biological membranes: Evolutionary conservation of an essential protein transport pathway (review). *Mol Membr Biol* 22:17–28.
2. Pool, M.R. (2005). Signal recognition particles in chloroplasts, bacteria, yeast and mammals (review). *Mol Membr Biol* 22:3–15.
3. Dalbey, R.E., and Chen, M. (2004). Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* 1694:37–53.
4. Van den Berg, B., Clemons, W.M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427:36–44.
5. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., III, Ban, N., and Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438:318–324.

6. Vrontou, E., and Economou, A. (2004). Structure and function of SecA, the preprotein translocase nanomotor. *Biochim Biophys Acta* 1694:67–80.
7. Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta* 1694:17–35.
8. Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W., von Heijne, G., van der Does, C., Driessen, A.J., Oudega, B., and Luirink, J. (2000). YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J* 19:542–549.
9. Bonnefoy, N., Chalvet, F., Hamel, P., Slonimski, P.P., and Dujardin, G. (1994). OXA1, a *Saccharomyces cerevisiae* nuclear gene whose sequence is conserved from prokaryotes to eukaryotes controls cytochrome oxidase biogenesis. *J Mol Biol* 239:201–212.
10. Bonnefoy, N., Kermorgant, M., Groudinsky, O., Minet, M., Slonimski, P.P., and Dujardin, G. (1994). Cloning of a human gene involved in cytochrome oxidase assembly by functional complementation of an oxa1-mutation in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 91:11978–11982.
11. Bauer, M., Behrens, M., Esser, K., Michaelis, G., and Pratje, E. (1994). PET1402, a nuclear gene required for proteolytic processing of cytochrome oxidase subunit 2 in yeast. *Mol Gen Genet* 245:272–278.
12. Altamura, N., Capitanio, N., Bonnefoy, N., Papa, S., and Dujardin, G. (1996). The *Saccharomyces cerevisiae* OXA1 gene is required for the correct assembly of cytochrome c oxidase and oligomycin-sensitive ATP synthase. *FEBS Lett* 382:111–115.
13. Kermorgant, M., Bonnefoy, N., and Dujardin, G. (1997). Oxa1p, which is required for cytochrome c oxidase and ATP synthase complex formation, is embedded in the mitochondrial inner membrane. *Curr Genet* 31:302–307.
14. He, S., and Fox, T.D. (1997). Membrane translocation of mitochondrially coded Cox2p: Distinct requirements for export of N and C termini and dependence on the conserved protein Oxa1p. *Mol Biol Cell* 8:1449–1460.
15. Hell, K., Herrmann, J., Pratje, E., Neupert, W., and Stuart, R.A. (1997). Oxa1p mediates the export of the N- and C-termini of pCoxII from the mitochondrial matrix to the intermembrane space. *FEBS Lett* 418:367–370.
16. Hell, K., Neupert, W., and Stuart, R.A. (2001). Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *EMBO J* 20:1281–1288.
17. Herrmann, J.M., Koll, H., Cook, R.A., Neupert, W., and Stuart, R.A. (1995). Topogenesis of cytochrome oxidase subunit II. Mechanisms of protein export from the mitochondrial matrix. *J Biol Chem* 270:27079–27086.
18. Stuart, R.A., and Neupert, W. (1996). Topogenesis of inner membrane proteins of mitochondria. *Trends Biochem Sci* 21:261–267.
19. Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and Stuart, R.A. (2003). Yeast Oxa1 interacts with mitochondrial ribosomes: the Importance of the C-terminal region of Oxa1. *EMBO J* 22:6438–6447.
20. Szyrach, G., Ott, M., Bonnefoy, N., Neupert, W., and Herrmann, J.M. (2003). Ribosome binding to the Oxa1 complex facilitates co-translational protein insertion in mitochondria. *EMBO J* 22:6448–6457.
21. Sundberg, E., Slagter, J.G., Fridborg, I., Cleary, S.P., Robinson, C., and Coupland, G. (1997). ALBINO3, an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. *Plant Cell* 9:717–730.
22. Moore, M., Harrison, M.S., Peterson, E.C., and Henry, R. (2000). Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. *J Biol Chem* 275:1529–1532.

23. Woolhead, C.A., Thompson, S.J., Moore, M., Tissier, C., Mant, A., Rodger, A., Henry, R., and Robinson, C. (2001). Distinct Albino3-dependent and -independent pathways for thylakoid membrane protein insertion. *J Biol Chem* 276:40841–40846.
24. Gerdes, L., Bals, T., Klostermann, E., Karl, M., Philippar, K., Hunken, M., Soll, J., and Schunemann, D. (2006). A second thylakoid membrane-localized Alb3/OxaI/YidC homologue is involved in proper chloroplast biogenesis in *Arabidopsis thaliana*. *J Biol Chem* 281:16632–16642.
25. Gohre, V., Ossenbuhl, F., Crevecoeur, M., Eichacker, L.A., and Rochaix, J.D. (2006). One of two alb3 proteins is essential for the assembly of the photosystems and for cell survival in *Chlamydomonas*. *Plant Cell* 18:1454–1466.
26. Bellafiore, S., Ferris, P., Naver, H., Gohre, V., and Rochaix, J.D. (2002). Loss of Albino3 leads to the specific depletion of the light-harvesting system. *Plant Cell* 14:2303–2314.
27. Ossenbuhl, F., Gohre, V., Meurer, J., Krieger-Liszkay, A., Rochaix, J.D., and Eichacker, L.A. (2004). Efficient assembly of photosystem II in *Chlamydomonas* reinhardtii requires Alb3. 1p, a homolog of Arabidopsis ALBINO3. *Plant Cell* 16:1790–1800.
28. Kuhn, A., Rohrer, J., and Gallusser, A. (1990). Bacteriophages M13 and Pf3 tell us how proteins insert into the membrane. *J Struct Biol* 104:38–43.
29. Wolfe, P.B., Rice, M., and Wickner, W. (1985). Effects of two sec genes on protein assembly into the plasma membrane of *Escherichia coli*. *J Biol Chem* 260:1836–1841.
30. Froderberg, L., Houben, E., Samuelson, J.C., Chen, M., Park, S.K., Phillips, G.J., Dalbey, R., Luirink, J., and De Gier, J.W. (2003). Versatility of inner membrane protein biogenesis in *Escherichia coli*. *Mol Microbiol* 47:1015–1027.
31. Rohrer, J., and Kuhn, A. (1990). The function of a leader peptide in translocating charged amino acyl residues across a membrane. *Science* 250:1418–1421.
32. Facey, S.J., and Kuhn, A. (2003). The sensor protein KdpD inserts into the *Escherichia coli* membrane independent of the Sec translocase and YidC. *Eur J Biochem* 270:1724–1734.
33. Bassilana, M., and Gwizdek, C. (1996). *In vivo* membrane assembly of the *E. coli* polytopic protein, melibiose permease, occurs via a Sec-independent process which requires the protonmotive force. *EMBO J* 15:5202–5208.
34. van Dalen, A., Schrempf, H., Killian, J.A., and de Kruijff, B. (2000). Efficient membrane assembly of the KcsA potassium channel in *Escherichia coli* requires the protonmotive force. *EMBO Rep* 1:340–346.
35. Yi, L., Jiang, F., Chen, M., Cain, B., Bolhuis, A., and Dalbey, R.E. (2003). YidC is strictly required for membrane insertion of subunits a and c of the F(1)F(0)ATP synthase and SecE of the SecYEG translocase. *Biochemistry* 42:10537–10544.
36. Van Der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A.J. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J Cell Biol* 165:213–222.
37. Yi, L., Celebi, N., Chen, M., and Dalbey, R.E. (2004). Sec/SRP requirements and energetics of membrane insertion of subunits a, b, and c of the *Escherichia coli* F1F0 ATP synthase. *J Biol Chem* 279:39260–39267.
38. van Bloois, E., Jan Haan, G., de Gier, J.W., Oudega, B., and Luirink, J. (2004). F(1)F(0) ATP synthase subunit c is targeted by the SRP to YidC in the *E. coli* inner membrane. *FEBS Lett* 576:97–100.
39. Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000). YidC mediates membrane protein insertion in bacteria. *Nature* 406:637–641.
40. Stuart, R.A., and Neupert, W. (2000). Making membranes in bacteria [comment] [news]. *Nature* 406:575–577.

41. Geller, B.L., and Wickner, W. (1985). M13 procoat inserts into liposomes in the absence of other membrane proteins. *J Biol Chem* 260:13281–13285.
42. Kiefer, D., and Kuhn, A. (1999). Hydrophobic forces drive spontaneous membrane insertion of the bacteriophage Pf3 coat protein without topological control. *EMBO J* 18:6299–6306.
43. van Dalen, A., van der Laan, M., Driessen, A.J., Killian, J.A., and de Kruijff, B. (2002). Components required for membrane assembly of newly synthesized K+ channel KcsA. *FEBS Lett* 511:51–58.
44. Ridder, A.N., Morein, S., Stam, J.G., Kuhn, A., de Kruijff, B., and Killian, J.A. (2000). Analysis of the role of interfacial tryptophan residues in controlling the topology of membrane proteins. *Biochemistry* 39:6521–6528.
45. Samuelson, J.C., Jiang, F., Yi, L., Chen, M., de Gier, J.W., Kuhn, A., and Dalbey, R.E. (2001). Function of YidC for the insertion of M13 procoat protein in *E. coli*: translocation of mutants that show differences in their membrane potential dependence and Sec-requirement. *J Biol Chem* 276:34847–34852.
46. Chen, M., Samuelson, J.C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R.E. (2002). Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion. *J Biol Chem* 277:7670–7675.
47. Froderberg, L., Houben, E.N., Baars, L., Luirink, J., and De Gier, J.W. (2004). Targeting and translocation of two lipoproteins in *Escherichia coli* via the SRP/Sec/YidC pathway. *J Biol Chem* 279:31026–31032.
48. Van Der Laan, M., Urbanus, M.L., Ten Hagen-Jongman, C.M., Nouwen, N., Oudega, B., Harms, N., Driessen, A.J., and Luirink, J. (2003). A conserved function of YidC in the biogenesis of respiratory chain complexes. *Proc Natl Acad Sci USA* 100:5801–5806.
49. Model, P., Jovanovic, G., and Dworkin, J. (1997). The *Escherichia coli* phage-shock-protein (psp) operon. *Mol Microbiol* 24:255–261.
50. Serek, J., Bauer-Manz, G., Struhalla, G., Van Den Berg, L., Kiefer, D., Dalbey, R., and Kuhn, A. (2004). *Escherichia coli* YidC is a membrane insertase for Sec-independent proteins. *EMBO J* 23:294–301.
51. Celebi, N., Yi, L., Facey, S.J., Kuhn, A., and Dalbey, R.E. (2006). Membrane biogenesis of subunit II of cytochrome bo oxidase: Contrasting requirements for insertion of N-terminal and C-terminal domains. *J Mol Biol* 357:1428–1436.
52. van Bloois, E., Haan, G.J., de Gier, J.W., Oudega, B., and Luirink, J. (2006). Distinct requirements for translocation of the N-tail and C-tail of the *Escherichia coli* inner membrane protein CyoA. *J Biol Chem* 281:10002–10009.
53. du Plessis, D.J., Nouwen, N., and Driessen, A.J. (2006). Subunit a of cytochrome o oxidase requires both YidC and SecYEG for membrane insertion. *J Biol Chem* 281:12248–12252.
54. Urbanus, M.L., Scotti, P.A., Froderberg, L., Saaf, A., de Gier, J.W., Brunner, J., Samuelson, J.C., Dalbey, R.E., Oudega, B., and Luirink, J. (2001). Sec-dependent membrane protein insertion: sequential interaction of nascent FtsQ with SecY and YidC. *EMBO Rep* 2:524–529.
55. Houben, E.N., Scotti, P.A., Valent, Q.A., Brunner, J., de Gier, J.L., Oudega, B., and Luirink, J. (2000). Nascent Lep inserts into the *Escherichia coli* inner membrane in the vicinity of YidC, SecY and SecA. *FEBS Lett* 476:229–233.
56. Beck, K., Eisner, G., Trescher, D., Dalbey, R.E., Brunner, J., and Muller, M. (2001). YidC, an assembly site for polytopic *Escherichia coli* membrane proteins located in immediate proximity to the SecYE translocon and lipids. *EMBO Rep* 2:709–714.
57. Mitra, K., and Frank, J. (2006). A model for co-translational translocation: ribosome-regulated nascent polypeptide translocation at the protein-conducting channel. *FEBS Lett* 580:3353–3360.

58. Bessonneau, P., Besson, V., Collinson, I., and Duong, F. (2002). The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *EMBO J* 21:995–1003.
59. Kaufmann, A., Manting, E.H., Veenendaal, A.K., Driessen, A.J., and van der Does, C. (1999). Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. *Biochemistry* 38:9115–9125.
60. Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004). Role of YidC in folding of polytopic membrane proteins. *J Cell Biol* 165:53–62.
61. Jiang, F., Yi, L., Moore, M., Chen, M., Rohl, T., Van Wijk, K.J., De Gier, J.W., Henry, R., and Dalbey, R.E. (2002). Chloroplast YidC homolog Albino3 can functionally complement the bacterial YidC depletion strain and promote membrane insertion of both bacterial and chloroplast thylakoid proteins. *J Biol Chem* 277:19281–19288.
62. van Bloois, E., Nagamori, S., Koningstein, G., Ullers, R.S., Preuss, M., Oudega, B., Harms, N., Kaback, H.R., Herrmann, J.M., and Luirink, J. (2005). The Sec-independent function of *Escherichia coli* YidC is evolutionary-conserved and essential. *J Biol Chem* 280:12996–13003.
63. Preuss, M., Ott, M., Funes, S., Luirink, J., and Herrmann, J.M. (2005). Evolution of mitochondrial oxa proteins from bacterial YidC. Inherited and acquired functions of a conserved protein insertion machinery. *J Biol Chem* 280:13004–13011.
64. Saaf, A., Monne, M., de Gier, J.W., and von Heijne, G. (1998). Membrane topology of the 60-kDa Oxa1p homologue from *Escherichia coli*. *J Biol Chem* 273:30415–30418.
65. Tjalsma, H., Bron, S., and van Dijl, J.M. (2003). Complementary impact of paralogous Oxa1-like proteins of Bacillus subtilis on post-translocational stages in protein secretion. *J Biol Chem* 278:15622–15632.
66. Yen, M.R., Harley, K.T., Tseng, Y.H., and Saier, M.H., Jr. (2001). Phylogenetic and structural analyses of the oxa1 family of protein translocases. *FEMS Microbiol Lett* 204:223–231.
67. Jiang, F., Chen, M., Yi, L., de Gier, J.W., Kuhn, A., and Dalbey, R.E. (2003). Defining the regions of *Escherichia coli* YidC that contribute to activity. *J Biol Chem* 278:48965–48972.
68. Nouwen, N., and Driessen, A.J. (2002). SecDFyajC forms a heterotetrameric complex with YidC. *Mol Microbiol* 44:1397–1405.
69. Nargang, F.E., Preuss, M., Neupert, W., and Herrmann, J.M. (2002). The Oxa1 protein forms a homooligomeric complex and is an essential part of the mitochondrial export translocase in Neurospora crassa. *J Biol Chem* 277:12846–12853.

This page intentionally left blank

# 5

# *Disulfide Bond Formation Enzymes*

JAMES C.A. BARDWELL

*Howard Hughes Medical Institute*
*University of Michigan*
*Ann Arbor, MI 48104, USA*

## I. Disulfides Stabilize Secreted Proteins

Proper protein function is dependent on the maintenance of pH and ionic concentrations in living organisms within very narrow ranges. The normal range for human blood pH is 7.4 $\pm$ 0.05, and the bacterium *Escherichia coli* is able to maintain its intracellular pH at 7.6 $\pm$ 0.2 even when exposed to extracellular pH ranges from 5.5 to 9 [1]. Proteins exposed to the harsh and often variable extracellular compartments, such as the bacterial periplasm, need an extra degree of stability in order to survive. To be able to tolerate these harsh and fluctuating conditions, many secreted proteins take advantage of the extra stability that disulfide bridges offer. In contrast, cysteines in the cytosolic space are generally maintained in the reduced configuration. This is in part due to the decreased need for protein stability and also because a number of classes of enzymes including cysteine proteases utilize reduced thiol groups in their catalytic mechanisms.

Catalytic systems have been discovered that maintain the thiol-disulfide-redox environment of these compartments within proper parameters. In particular, disulfide bond formation and isomerization are catalyzed processes in both prokaryotes and eukaryotic organisms. This area has been the subject of a number of reviews [2–4]. This chapter will discuss disulfide bond formation and isomerization in the periplasm of the model organism *E. coli*. In keeping with the title of this series, this chapter will focus on the enzymatic properties of the disulfide catalysts and isomerases.



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25005-X

## II. The Need for a Catalyst

A general scheme for the formation of a disulfide is as follows:

$$R—SH + R'—SH \rightarrow R—S—S—R' + 2e^- + 2H^+ \quad (1)$$

The reaction shown is an oxidation reaction with respect to the thiols, the reverse reaction is a reduction reaction. The oxidation of thiols requires an appropriate electron acceptor, and their reduction requires an electron donor. Like many biologically important reactions, disulfide bond formation can take place spontaneously. Disulfide bonds can be formed in the presence of oxygen according to the following formula:

$$2R—SH + \frac{1}{2}O_2 \rightarrow RS—SR + H_2O \quad (2)$$

However, as for most biological reactions, the uncatalyzed rate of this reaction is very slow and inadequate to sustain life. Redox active metals, such as iron and copper, can act via the Fenton reaction to generate reactive oxygen species that in turn act to form disulfides; but under most circumstances, the concentration of these metals in the cell is vanishingly small [5]. Thus, there would appear a need for a catalyst to form these bonds. In *E. coli*, disulfide bonds are formed in the periplasmic space. The enzymes responsible are called the disulfide bond (Dsb) enzymes for their ability to affect the formation and isomerization of disulfide bonds [2]. DsbA and DsbB are responsible for thiol oxidation and DsbC, DsbG, and DsbD are responsible for disulfide isomerization.

## III. DsbA: The Primary Oxidant

DsbA is a 21-kDa soluble periplasmic protein that is part of the thioredoxin family. It is thought to serve as the primary source of disulfide bonds to secreted proteins [6]. A disulfide formed between the Cys-X-X-Cys (CXXC; where X represents any amino acid) motif present at its active site is directly donated to secreted proteins catalyzing their oxidation and thus their folding (Figure 5.1).

Disulfide transfers from DsbA to substrate protein. Oxidized DsbA rapidly reacts with unfolded proteins entering the periplasm. The disulfide is transferred from DsbA to the protein, resulting in the oxidation of the substrate and the reduction of DsbA.

The reaction catalyzed by DsbA is a very simple thiol–disulfide exchange. A number of factors discussed further make DsbA well suited

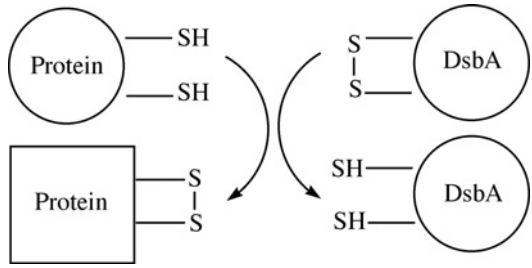



FIG. 5.1. The oxidation of a substrate protein by DsbA. DsbA donates it active site disulfide to a substrate protein via a thiol–disulfide exchange reaction.

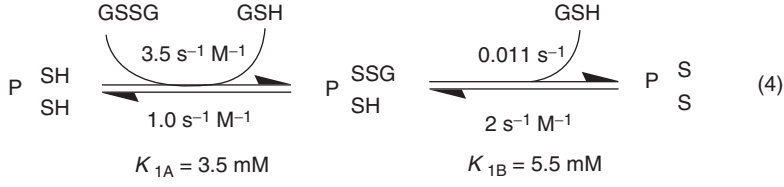
to perform this reaction, but it is first worth noting that though null mutants in *dsbA* have substantially decreased rates of disulfide bond formation, the rate in these mutants is still measurable and can be physiologically relevant. The protein OmpA for instance is nearly fully oxidized within a 40-s pulse period in wild-type strains but is nearly 50% oxidized after 5 min in *dsbA* null mutations [6]. This represents a substantial enhancement in the rate of disulfide bond formation by DsbA. However, non-DsbA-dependent disulfide bond formation still occurs at adequate rates to allow for the accumulation of near wild-type quantities of oxidized OmpA and many other disulfide-containing proteins in *dsbA* null mutant bacteria [7]. We have observed that many of the phenotypes of *dsbA* mutants can be at least partially compensated for by exogenous addition of small disulfide-containing molecules, such as cystine or oxidized glutathione or redox active metals such as copper. Examining these alternative sources of disulfide bonds in the cell gives us a perspective on whether DsbA should be considered an enzyme or if it is also realistic to consider its properties as a redox buffer. Most thiol–disulfide exchange reactions follow the simple formula:

$$R_1S — + \ R_2SSR_3 \rightarrow R_1SS_3R + R_2S- \tag{3}$$

where the thiolate anion $R_1S-$ displaces one sulfur of the disulfide bond.

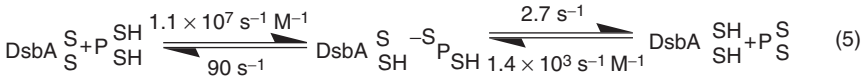
This is the case whether the R groups are part of proteins or small molecules. Disulfide bonds in proteins are formed by two subsequent thiol–disulfide exchange reactions with the redox reagent. The first results in a mixed disulfide between the protein and the redox reagent. The reaction of glutathione with proteins, like the reaction with DsbA with its protein substrates, also results in the oxidation of the proteins and the reduction of the glutathione. Darby and Creighton used a two-cysteine-containing peptide that adopts only local nonrandom conformations as

a simple model of an unfolded protein [8]. The direct comparison performed by Darby and Creighton of the reactions of DsbA and oxidized glutathione (GSSG) with this peptide reveals at least two reasons why it is generally accepted that DsbA is an enzyme.



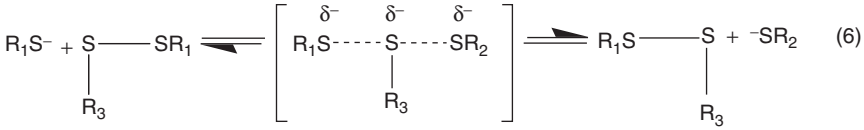

(4)

$$K_1 = K_{1A}K_{1B} = 1.9 \times 10^{-2}\ \mathrm{M}$$




(5)

First, DsbA is capable of much more rapid thiol–disulfide exchange with proteins than glutathione is; in general, thiol–disulfide exchange reactions involving DsbA occur $10^2$–$10^6$ times more rapidly than is typically observed for thiols and disulfides [8]. Furthermore, the overall equilibrium of the reaction is very strongly shifted in the direction of protein oxidation. Together these factors combine to result in the observation that DsbA oxidizes this unfolded protein model proteins much faster than glutathione does. Thus, DsbA is a very strongly oxidizing protein, one of the most strongly oxidizing if not the most strongly oxidizing thiol–disulfide oxidoreductase known. The strongly oxidizing nature of DsbA and its capacity to undergo very rapid thiol–disulfide exchange reactions allows DsbA to rapidly oxidize proteins *in vivo* [9]. Much of DsbA's catalytic mechanism can be understood in chemical terms by the factors influencing thiol–disulfide exchange reactions between small molecules. The disulfide interchange reaction diagrammed in Eq. 6 is a concerted reaction going through a single transition state.




(6)

Thus, during thiol–disulfide exchange, an intermediate is formed where partial negative charge is distributed over all three sulfur atoms. The sulfur

atom that accepts the electron is the one that is the most electrophilic. A measure of the electrophilicity is the p$K_a$ of the thiol groups, this determines not only their intrinsic reactivity but also their reactivity when present within a disulfide bond. For DsbA, the p$K_a$ of the N-terminal cysteine of the active site is about 3.5 [10]. This is an unusually low value, the p$K_a$ cysteine free in solution or in disordered peptides are around 9. The p$K_a$ is the pH at which 50% of a species is in an ionized form. This low p$K_a$ ensures that DsbA will remain ionized and thus capable of undergoing disulfide exchange reactions even when the cell is exposed to acidic conditions. Thus, the rate of exchange observed at pH 4 between DsbA and a model peptide are comparable with those observed at pH 7.4. In contrast, thiol–disulfide exchange reactions between the peptide and glutathione are negligible at pH 4 [8, 10].

More importantly, this low p$K_a$ helps explain the strongly oxidizing redox potential of DsbA in both a qualitative and quantitative manner. Qualitatively, when thiol groups are ionized, they cannot be part of a disulfide bond. The strong stabilization of the Cys30 thiolate ion results in the unusual situation where DsbA as a protein is actually more stable in the reduced than in the oxidized form [9]. Disulfides generally stabilize protein folds. The stabilization of the reduced form provides the thermodynamic driving power behind the donation of DsbA's disulfide to other proteins. Quantitatively, it has been shown for small molecules that one can predict the rate constant of a thiol exchange reaction if one knows the p$K_a$s of all three thiol groups involved using the following formula:

$$\log k = 4.5 + 0.59\,\mathrm{p}K_{a1} - 0.4\,\mathrm{p}K_{a2} - 0.59\,\mathrm{p}K_{a3} \tag{7}$$

The p$K_a$ of Cys30 of DsbA is around 3.5 [10]. The p$K_a$ of Cys33 is around 14 (Kreissig, Zander, and Bardwell, unpublished data). The p$K_a$ of cysteines in unfolded proteins is very similar to that of glutathione, that is, around 8.7 [10]. From the microscopic rate constants of the reactions of DsbA with glutathione, one can calculate the overall equilibrium constant of DsbA with glutathione. This equilibrium constant is also known as $K_{ox}$ and is a standard measure of the oxidizing power of proteins [11].

$$\mathrm{DsbA_{red}} + \mathrm{GSSG} \rightleftharpoons \mathrm{DsbA_{ox}} + 2\mathrm{GSH} \tag{8}$$

$$K_{ox} = \frac{[\mathrm{DsbA_{ox}}]\ [\mathrm{GSH}]^2}{[\mathrm{DsbA_{red}}]\ [\mathrm{GSSG}]} \tag{9}$$

The $K_{ox}$ calculated for DsbA using these p$K_a$ values is 0.26 mM similar to the value of 0.12 mM, which was measured experimentally with DsbA

[12]. Thus, the strongly oxidizing redox potential of DsbA and its ability to undergo rapid thiol exchange reactions are in large part explained by the unusual p$K_a$ of DsbA's cysteines. These unusual p$K_a$ values in turn are explained almost entirely by the precise positioning of the residues near the active site of DsbA. An empirical method called PROPKA has been developed that allows one to calculate the p$K_a$ values of all ionizable groups on any protein whose structure is known [13]. When applied to DsbA, this method accurately predicts the p$K_a$ of the N- and C-terminal cysteines in DsbA to be 3.4 and 15, respectively. This methodology suggests that three residues, Cys33, His32, and Ser27, are primarily responsible for the unusually low p$K_a$ of Cys30, due to hydrogen bond interactions with main chain atoms [13]. It has been experimentally shown that His32 is partially responsible for the low p$K_a$ of Cys30, mutations that change residue 32 in DsbA affect its redox potential in ways that at least qualitatively can be explained by their effect on p$K_a$ values [12, 14]. The Cys30 thiol in Cys33Ser substitutions of DsbA is, however, still substantially ionized at pH 5 suggesting that at least for this substitution the p$K_a$ of Cys30 is not substantially altered [10]. Nelson and Creighton have also suggested that the positively charged dipole at the N-terminus of the α-helix on which Cys30 is situated in DsbA may be at least partially responsible for lowering the p$K_a$ of Cys30 [10]. In summary, it appears that the strongly oxidizing redox potential of DsbA and its ability to undergo rapid thiol–disulfide exchange reactions can be understood at least in part in electrostatic terms.

## IV. Structure of DsbA

High-resolution structures are available for both oxidized and reduced *E. coli* DsbA and for various mutants of *E. coli* DsbA and DsbA from other species [15–19]. DsbA has a thioredoxin-like fold. This fold is remarkably common in thiol–disulfide oxidoreductases. For example, it is present in four of the five Dsb enzymes found in prokaryotes namely DsbA, DsbC, DsbD, and DsbG. In general, the thioredoxin-like fold is a distinct structural motif minimally composed of a four-stranded β-sheet and three flanking α-helices [20]. At the end of the α1 helix of the thioredoxin fold in DsbA, thioredoxin and glutaredoxin is a CXXC motif that reversibly forms a disulfide. This is the disulfide that DsbA donates to other proteins. The various properties of the different oxidoreductases of the thioredoxin family can be understood in terms of how their different structures modify their redox potentials, substrate specificities, and localizations (Pan and Bardwell, [73]). DsbA in addition, contains an α-helical domain, which is inserted into the thioredoxin fold at a point between β2 and α2.

This α-helical insertion contains several hydrophobic residues that may contribute to the peptide-binding properties of DsbA. Several observations indicate that DsbA interacts with peptides in a noncovalent fashion. Couprie showed by NMR that DsbA binds model peptides in a hydrophobic manner; hydrophobic interactions are a common theme among protein-folding chaperones and catalysts [21].

Second, DsbA reacts with peptides 300- to 1300-fold faster than it does with glutathione, the larger the peptide the faster the reaction, further indications that DsbA interacts noncovalently with peptides [8]. Schmid and coworkers have shown that in a mixed disulfide between DsbA and RnaseT1 the conformation of T1 is decreased by about 5 kJ $mol^{-1}$, while the stability of DsbA is increased by about the same amount, suggesting that DsbA interacts noncovalently with folding proteins [22]. However, as previously noted [8], the DsbA/peptide nonconvalent complex need not be very stable to produce these observed rate enhancements and changes in stability. In general, the peptide-binding properties of DsbA remain much more poorly characterized than its redox properties.

## V. How Is DsbA Reoxidized?

Enzymes must return to their original state following a round of catalysis. However, after DsbA has oxidized its substrate protein, DsbA is present in a reduced form and is incapable of oxidizing another protein. In order for it to be catalytic, DsbA needs to be reoxidized. The membrane-bound protein DsbB performs this role. Mutations in *dsbB* were originally identified using a selection for cells that were deficient in efficiently forming disulfides, the same selection that was used to find *dsbA* mutations [23]. Like *dsbA* mutants, *dsbB* mutants are pleiotropically defective in forming disulfides in the cell [23–25].

DsbB is an integral membrane protein with four transmembrane helices and two periplasmic domains. Each of these periplasmic loops contains a pair of cysteines that are essential for DsbB activity [26]. These cysteines are involved in a disulfide catalytic cascade reaction that ends in the transfer of a disulfide to DsbA. Although the exact sequence of the disulfide cascade is still unclear, most of the experimental evidence points to the cysteines in the more C-terminal domain serving as the immediate disulfide donor to DsbA [2]. If one prevents the completion of transfer of disulfides to DsbA by eliminating the C-terminal cysteine in DsbA, a mixed disulfide between Cys30 of DsbA and Cys104 of DsbB can be isolated, supporting this model [27]. The disulfide in the N-terminal periplasmic domain of DsbB that forms between Cys41 and Cys44 is then thought to be transferred to the C-terminal periplasmic domain to form

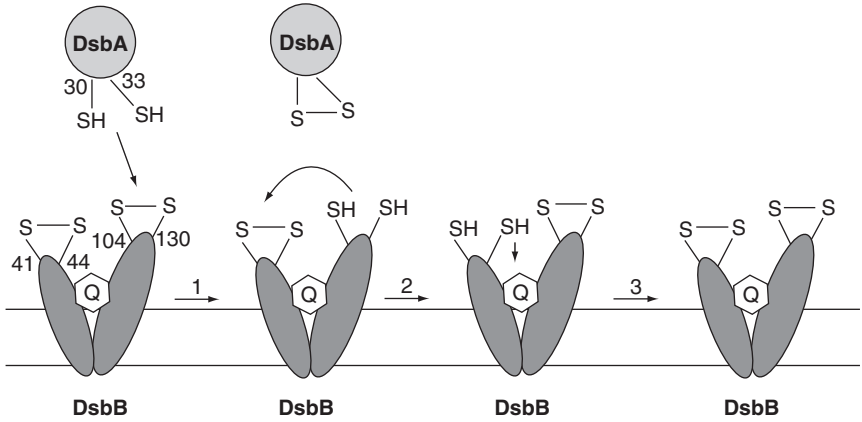



FIG. 5.2. Reoxidation of DsbA by DsbB. Reduced DsbA is reoxidized by the Cys104-Cys130 disulfide of DsbB. Arrows show the direction of electron flow. Electrons are transferred from DsbA to the Cys104-Cys130 disulfide of DsbB. This is followed by an electron transport cascade within DsbB where a bound quinone serves as the final electron acceptor within DsbB.

Cys104-Cys130. Finally, electrons are donated to a bound quinone, regenerating fully oxidized DsbB. Other more complicated models have also been proposed [27] (Figure 5.2).

This model is only partially supported by measurements of the redox potentials of DsbB and its thiol–disulfide pairs. When equimolar quantities of oxidized DsbB and reduced DsbA are mixed and allowed to reach equilibrium, ~40% of the DsbA becomes oxidized, suggesting that the overall redox potential of DsbB is similar to that of DsbA [28]. However, all but one of the measurements of the redox potential of any of the individual cysteine pairs in DsbB come out to be less oxidizing than DsbA. The redox potential of the Cys104-Cys130 disulfide in DsbB has been measured to be either −250, −284, or −186 mV, in any case a substantially lower redox potential than that of Cys30-Cys33 in DsbA (−120 mV) [29–31]. In general, electrons flow down an energy gradient, thus DsbA should be able to oxidize the Cys104-Cys130 disulfide, which is the reverse of the physiological direction; indeed, when oxidized DsbA is added to reduced DsbB which contains only two cysteines Cys104 and Cys130, DsbA is reduced and DsbB is oxidized [30]. The Cys40-Cys44 pair has been determined to have a redox potential of −270 [30, 32], −210 [29], or −69 mV [31]. Of these, only the −69-mV value is more oxidizing than the redox potential of DsbA, but this value appears to be an artifact of the use of fluorescence changes in DsbB to measure its redox status. Inaba *et al.* [28] concluded that these fluorescence intensity changes do not report changes in the DsbB protein itself and are thus not appropriate to

measure redox potentials of DsbB's cysteine pairs [28]. Thus, by most measures, the redox potential of the individual disulfide pairs in DsbB is less oxidizing than that of DsbA. How then can DsbB reoxidize DsbA? One potential solution to this dilemma is to postulate that the redox potential of DsbB is increased by DsbA binding or that the redox potential of the two pairs of cysteines in DsbB affects each other. Another way out is to suppose that the reoxidation of DsbB drives the disulfide cascade reaction within DsbB.

## VI. Reoxidation of DsbB

Each time a disulfide is formed, two electrons are generated. Thus, the search for the reoxidant of DsbB is a search for an electron acceptor. DsbA is reduced during growth under deprivation of quinones. A *ubiA-menA* double mutant strain that is defective in the synthesis of ubiquinone and menaquinone leads to the accumulation of reduced DsbA, which implies that quinones are involved in the reoxidation of DsbA [33, 34]. We were able to reconstitute the entire disulfide bond catalytic system *in vitro* [35]. We showed that electrons flow from DsbB to ubiquinone, then on to cytochrome oxidases *bd* and *bo*, and finally on to molecular oxygen.

Cytochrome *bo* oxidase is abundant under aerobic conditions [36]. Cytochrome *bd* oxidase is expressed in response to limiting oxygen conditions and has a $K_m$ for oxygen estimated to be as low as 20 nM; it may thus be able to scavenge enough oxygen to allow disulfide bond formation to proceed even under very low oxygen conditions [35, 37]. Under truly anaerobic conditions, DsbB donates its electrons to menaquinone which in turn is reoxidized by anaerobic electron acceptors such as fumarate, dimethyl sulfoxide (DMSO), or nitrate reductase [35, 38]. By alternating electron acceptors depending on the growth conditions, disulfide bond formation can thus take place during growth under various oxygen tensions. DsbB thus is a central player in disulfide bond formation in prokaryotes as it links disulfide bond formation to electron transport. Reactions downstream from DsbB involve thiol–disulfide exchange reactions; reactions upstream from DsbB involve electron transport. DsbB uses the oxidizing power of quinones to generate disulfides *de novo* [34, 35]. Thus, DsbB acts as a quinone reductase [39]. The redox potential of ubiquinone is approximately +100 mV; thus, the oxidizing power of quinones provides the thermodynamic power driving the overall reaction.

DsbB undergoes a striking color transition during its reaction with DsbA. This purple color represents an intermediate in the reaction mechanism [32, 40]. This color is likely due to a novel quinone charge transfer complex [32]. With bound ubiquinone, the reaction intermediate has a strong purple

absorption at 500 nm, with bound menaquinone, DsbB develops a blue color ($\lambda_{max} = 550$ nm) during its reaction [38]. These colors provide a convenient way to monitor the reaction between DsbB and quinones. One possibility is that the colored species is a thiolate-quinone charge transfer complex. The purple color is dependent on the presence of the Cys44, presumably as thiolate ion because the purple color can be observed whenever Cys44 is not part of a disulfide bond [40]. Mutations that replace an adjacent highly conserved arginine Arg48 abolish the purple color [34]. They also decrease the amount of bound ubiquinone (UQ), and increasing DsbB's $K_m$ for UQ sevenfold [34,41]. It has been proposed that the role of Arg48 is to stabilize a charge transfer complex by electrostatically counteracting the Cys44 thiolate ion [34].

Alternatively, it has been proposed that the purple color is due to a quinhydrone [32]. A quinhydrone is a stacked configuration of two quinones, one oxidized, one reduced. In this model, one quinone is very tightly bound to DsbB (resident quinone), while the second quinone is interchangeable. Electrons are transferred from DsbB to the resident quinone and from there to the second quinone. This quinone becomes reduced and transfers the electrons to components of the respiratory chain. To start a new cycle, oxidized quinone from the quinone pool replaces the reduced quinine (Figure 5.3).

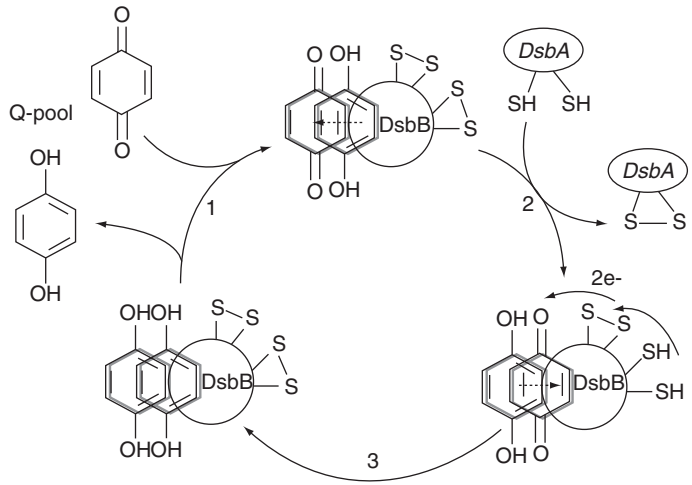



FIG. 5.3. One model for the reaction cycle of the quinone reductase activity of DsbB. The quinone reductase activity of DsbB involves two quinines: a prosthetic quinone that remains bound to DsbA and a more transient quinone. Interaction between the prosthetic quinone and the transient quinone allows the transfer of electrons from DsbB to the free quinone. This allows the oxidation of DsbB and in turn the oxidation of DsbA. Another model involves a single oxidized quinone binding to DsbB and then being reduced by DsbB followed by quinone dissociation.

If oxidized quinone is added to the quinone-free DsbB (SCSS) mutant, the protein solution turns purple, and this purple color persists [40]. Since there is no source of reduced quinone in these experiments, this provides strong evidence that this stable purple color is not due to a quinhydrone. However, one quinone is so strongly bound to DsbB that it can be purified with it. The pool of free quinone present in the membrane reoxidizes this bound quinone. The most straightforward way for this reoxidation to occur is via direct interaction of these two quinones by the at least transiently formation of a quinhydrone charge transfer complex. This species may contribute to the color transitions seen in DsbB's reaction cycle.

## VII. Disulfide Bond Isomerization

DsbA is a strong oxidant and is capable of very rapid thiol–disulfide exchange reactions. If a protein only has two-cysteine residues that form a single disulfide bond, then the only possible way a disulfide can be formed is the correct one and even very powerful oxidant will get it right. However, as the number of cysteine residues in a protein increases above 2, the number of theoretically possible intermolecular disulfides increases very rapidly, and with it, the possibility of nonnative oxidation. One consequence of the strongly oxidizing rapid nature of DsbA is that it is likely to oxidize cysteines on proteins very soon after they are secreted into the periplasm. If the native disulfides link cysteines that are consecutive in the sequence (formed between cysteine residues that are nearest to each other in the primary sequence), then the correct disulfide linkages will be formed even by a powerful oxidant such as DsbA. If, however, the native disulfides require linking nonconsecutive cysteine residues, then there exists a very real possibility that nonnative linkages occur; this will disrupt the structure of the protein and lead to its inactivation. In order to correct these nonnative disulfide linkages, cells have evolved disulfide isomerases. In *E. coli*, two proteins with significant disulfide isomerase activity have been isolated, DsbC and DsbG. These two proteins share 24% sequence identity, a common structure, and many similar properties. They appear to differ however in their *in vivo* role and substrate specificity.

DsbC can rearrange disulfides *in vivo* and *in vitro* [42–44]. For example, addition of catalytic quantities of DsbC to bifunctional inhibitor from Ragi (RBI) (a protein with five nonconsecutive disulfides) leads to its functional refolding [45]. It is much more active in isomerization than DsbA is and seems particularly good at rearranging buried, nonnative disulfides that have been formed by the oxidant GSSG. *In vivo* DsbC expression or

overexpression is either required or helpful for the efficient expression of a number of recombinant proteins with nonconsecutive disulfides, including bovine pancreatic trypsin inhibitor (BPTI), urokinase, tPA, insulin-like growth factor I [46, 47], aporetinol-binding protein, horseradish peroxidase, brain-derived neurotropic factor [48–51] and single-chain Fv antibody fragments [52], and melanocyte growth-stimulating activity. DsbG when overexpressed is capable of at least partially restoring the ability of DsbC mutants to express functionally multidisulfide proteins such as BPTI and mouse urokinase [53]. DsbC and DsbG also have protein chaperone activity, which may assist in their ability to fold periplasmic proteins [54, 55]. DsbC is required for the expression of a number of native *E. coli* proteins that contain at least one nonconsecutive disulfide, including RNaseI and phytase [56, 57]. In contrast, there are no known *E. coli* substrates for DsbG. It may be that it has much more limited substrate specificity than DsbC does; it is also possible that it plays a role *in vivo* unrelated to its isomerase activity.

There are at least two possible mechanisms whereby DsbC can isomerase disulfides in proteins. Both mechanisms involve the attack of a nonnative disulfide in a substrate protein by reduced DsbC, forming a mixed disulfide between the protein and DsbC. These mixed disulfides accumulate during DsbC-assisted folding of RBI [45]. A second cysteine in the substrate protein can attack this mixed disulfide resulting in a native disulfide in the substrate and reduced DsbC. Alternatively, the second cysteine in DsbC may attack the mixed disulfide. This will result in the transfer of the disulfide to DsbC's active site and the reduction of the nonnative disulfide. This may then be followed by a second attempt at correct oxidation of the substrate either by DsbA or possibly by oxidized DsbC.

DsbC is active in attacking incorrect disulfides only if it is kept in the reduced form. This is done by the protein DsbD. Curiously the redox potential of DsbC (−135 mV) is almost as oxidizing as that of DsbA (−122 mV) [9, 58, 59]. How then can DsbC possibly attack incorrect disulfides? One possibility is that incorrect disulfides are in general so strained that they have redox potentials even more oxidizing than that of DsbC. The overall redox potential of an isomerase is more important if the isomerization mechanism involves complete reduction of the incorrect disulfide by DsbC. The rate of the attack of incorrect disulfides in substrates to form mixed disulfides with the catalyst is one of the key reactions that determine if a protein will be an effective isomerase. Darby *et al.* [44] using a model peptide found that this attack occurs 6000 times faster with DsbC than it does with DsbA [44].

DsbC and DsbG, like DsbA, are built around a thioredoxin-like fold, but they appear to have diverged in the distant past and they share only about 10% sequence identity. DsbC and DsbG lack the α-helical domain

that is inserted into DsbA relative to thioredoxin between β2 and α2. Instead DsbC and DsbG have a dimerization domain added on to the N-terminus of the thioredoxin fold. These isomerases form V-shaped dimers with the thioredoxin-like folds facing the interior of the dimer [60, 61]. Eukaryotic protein disulfide shares this overall architecture, and there is evidence that two CXXC active site motifs are required for isomerase activity [62]. Heterodimers of DsbC where one of the active sites are inactivated completely lack isomerase activity but are unchanged in their oxidation activity and the dimeric form of DsbC is important for both its chaperone and isomerase activity [48]. One can even bestow isomerase activity on DsbA, thioredoxin, and the thioredoxin-like domain of PDI by fusing them to the dimerization domain of DsbC [57, 63]. Mutants of DsbC that disrupt its ability to dimerize are substrates for DsbB and function as oxidases [64]. It appears that the proper orientation of the active sites controls these chimera's function as oxidases or isomerases. The increase in local concentration of redox active cysteines may be one of the key features that allow these dimeric proteins to function as isomerases. One of the active sites can be involved in reduction reactions, while the other is involved in oxidation or isomerization reactions with multiple cysteines on the same substrate protein.

## VIII. DsbD a Disulfide Transporter?

DsbD is an integral membrane protein whose function is to keep DsbC in a reduced state. This is a prerequisite for DsbC's ability to function as an isomerase [25, 42, 65]. DsbD consists of three discrete domains, two periplasmic domains, the α- and the γ-domain sandwiching a single-membrane bound domain, which is referred to as the β-domain. Each of the three domains possesses a pair of cysteines that are essential for the function of DsbD [66]. These cysteines appear to be involved in a disulfide cascade mechanism that results in the reduction of the cysteines in DsbC on the periplasmic side of the membrane and the oxidation of the cysteines in thioredoxin which is located on the cytoplasm side of the domain [66, 67]. Thus, the overall function of DsbD is to shuttle reducing equivalents from the cytoplasmic protein thioredoxin to the periplasmic protein DsbC. It is thought that electrons flow sequentially from thioredoxin to β, to γ, to α, and finally on to the isomerases DsbC and DsbG [68] (Figure 5.4).

This reaction order is consistent with the measured redox potentials of the interacting partners. In order from least to most oxidizing comes thioredoxin with −270, the γ-domain with −241, the α-domain with −229, and DsbC with −135 mV [69]. The three domains of DsbD can be

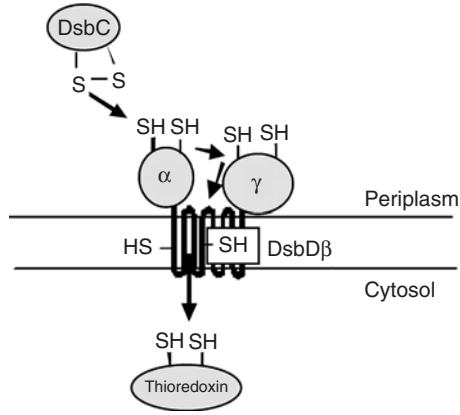



FIG. 5.4. Isomerization pathway in the *E. coli* periplasm. In order for DsbC to function as an isomerase, it must be in the reduced form. It is kept this way by the membrane protein DsbD. The proposed direction of disulfide flow is shown by arrows. DsbD is kept reduced by the reducing power of thioredoxin.

expressed separately from individual plasmids *in vivo* to form a functional unity in the cell [68]. Using this approach, the authors could show that when the immediate upstream electron donor was missing, the remaining domains accumulated in the oxidized form. This shows that the electron transfer is conducted in a sequential manner of oxidation/reduction reactions of the individual domains. The predicted intermediates between DsbDα and DsbC and DsbDβ and thioredoxin could be isolated [70, 71]. The simplest model of electron transport within DsbD involves a disulfide cascade reaction. However, getting the disulfide bonds across the inner membrane poses an interesting topological puzzle: how do you transport a disulfide bond across a membrane? The mechanism of how the disulfide of DsbDβ alternatively interacts with the cysteines in DsbD$\gamma$ and the cysteines of thioredoxin is still unclear. The disulfide of DsbDβ is located at the cytoplasmic site and the cysteines of DsbD$\gamma$ is in the periplasm. DsbDβ cysteines *either* must have access to both sides of the membrane simultaneously *or* must go from a conformation where they first have access to the periplasm, and then to one with access to the cytoplasm. The latter is likely to require a significant conformational change [72]. Structural characterization of the full-length protein and/or intermediates of the electron transfer should give further insight into the detailed molecular mechanism.

## References


1. Slonczewski, J.L., Rosen, B.P., Alger, J.R., and Macnab, R.M. (1981). pH homeostasis in *Escherichia coli*: measurement by 31P nuclear magnetic resonance of methylphosphonate and phosphate. *Proc Natl Acad Sci USA* 78:6271–6275.
2. Kadokura, H., Katzen, F., and Beckwith, J. (2003). Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem* 72:111–135.
3. Nakamoto, H., and Bardwell, J.C. (2004). Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta* 1694:111–119.
4. Collet, J.F., and Bardwell, J.C. (2002). Oxidative protein folding in bacteria. *Mol Microbiol* 44:1–8.
5. Lee, J., Pena, M.M., Nose, Y., and Thiele, D.J. (2002). Biochemical characterization of the human copper transporter Ctr1. *J Biol Chem* 277:4380–4387.
6. Bardwell, J.C., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulfide bond formation *in vivo*. *Cell* 67:581–589.
7. Leichert, L.I., and Jakob, U. (2004). Protein thiol modifications visualized *in vivo*. *PLoS Biol* 2:e333.
8. Darby, N.J., and Creighton, T.E. (1995). Catalytic mechanism of DsbA and its comparison with that of protein disulfide isomerase. *Biochemistry* 34:3576–3587.
9. Zapun, A., Bardwell, J.C., and Creighton, T.E. (1993). The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation *in vivo*. *Biochemistry* 32:5083–5092.
10. Nelson, J.W., and Creighton, T.E. (1994). Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation *in vivo*. *Biochemistry* 33:5974–5983.
11. Gilbert, H.F. (1990). Molecular and cellular aspects of thiol-disulfide exchange. *Adv Enzymol Relat Areas Mol Biol* 63:69–172.
12. Guddat, L.W., Bardwell, J.C., Glockshuber, R., Huber-Wunderlich, M., Zander, T., and Martin, J.L. (1997). Structural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci* 6:1893–1900.
13. Li, H., Robertson, A.D., and Jensen, J.H. (2005). Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 61:704–721.
14. Grauschopf, U., Winther, J.R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J.C. (1995). Why is DsbA such an oxidizing disulfide catalyst? *Cell* 83:947–955.
15. Martin, J.L., Bardwell, J.C., and Kuriyan, J. (1993). Crystal structure of the DsbA protein required for disulphide bond formation *in vivo*. *Nature* 365:464–468.
16. Guddat, L.W., Bardwell, J.C., Zander, T., and Martin, J.L. (1997). The uncharged surface features surrounding the active site of *Escherichia coli* DsbA are conserved and are implicated in peptide binding. *Protein Sci* 6:1148–1156.
17. Guddat, L.W., Bardwell, J.C., and Martin, J.L. (1998). Crystal structures of reduced and oxidized DsbA: investigation of domain motion and thiolate stabilization. *Structure* 6:757–767.
18. Schirra, H.J., Renner, C., Czisch, M., Huber-Wunderlich, M., Holak, T.A., and Glockshuber, R. (1998). Structure of reduced DsbA from *Escherichia coli* in solution. *Biochemistry* 37:6263–6276.
19. Hu, S.H., Peek, J.A., Rattigan, E., Taylor, R.K., and Martin, J.L. (1997). Structure of TcpG, the DsbA protein folding catalyst from Vibrio cholerae. *J Mol Biol* 268:137–146.
20. Martin, J.L. (1995). Thioredoxin—a fold for all reasons. *Structure* 3:245–250.

21. Couprie, J., Vinci, F., Dugave, C., Quéméneur, E., and Moutiez, M. (2000). Investigation of the DsbA mechanism through the synthesis and analysis of an irreversible enzyme-ligand complex. *Biochemistry* 39:6732–6742.
22. Frech, C., Wunderlich, M., Glockshuber, R., and Schmid, F.X. (1996). Preferential binding of an unfolded protein to DsbA. *EMBO J* 15:392–398.
23. Bardwell, J.C., Lee, J.O., Jander, G., Martin, N., Belin, D., and Beckwith, J. (1993). A pathway for disulfide bond formation *in vivo*. *Proc Natl Acad Sci USA* 90:1038–1042.
24. Dailey, F.E., and Berg, H.C. (1993). Mutants in disulfide bond formation that disrupt flagellar assembly in *Escherichia coli*. *Proc Natl Acad Sci USA* 90:1043–1047.
25. Missiakas, D., Georgopoulos, C., and Raina, S. (1993). Identification and characterization of the *Escherichia coli* gene dsbB, whose product is involved in the formation of disulfide bonds *in vivo*. *Proc Natl Acad Sci USA* 90:7084–7088.
26. Jander, G., Martin, N.L., and Beckwith, J. (1994). Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. *EMBO J* 13:5121–5127.
27. Kadokura, H., and Beckwith, J. (2002). Four cysteines of the membrane protein DsbB act in concert to oxidize its substrate DsbA. *EMBO J* 21:2354–2363.
28. Inaba, K., Takahashi, Y.H., and Ito, K. (2005). Reactivities of quinone-free DsbB from *Escherichia coli*. *J Biol Chem* 280:33035–33044.
29. Inaba, K., and Ito, K. (2002). Paradoxical redox properties of DsbB and DsbA in the protein disulfide-introducing reaction cascade. *EMBO J* 21:2646–2654.
30. Regeimbal, J., and Bardwell, J.C. (2002). DsbB catalyzes disulfide bond formation *de novo*. *J Biol Chem* 277:32706–32713.
31. Grauschopf, U., Fritz, A., and Glockshuber, R. (2003). Mechanism of the electron transfer catalyst DsbB from *Escherichia coli*. *EMBO J* 22:3503–3513.
32. Regeimbal, J., Gleiter, S., Trumpower, B.L., Yu, C.A., Diwakar, M., Ballou, D.P., and Bardwell, J.C. (2003). Disulfide bond formation involves a quinhydrone-type charge-transfer complex. *Proc Natl Acad Sci USA* 100:13779–13784.
33. Kobayashi, T., Kishigami, S., Sone, M., Inokuchi, H., Mogi, T., and Ito, K. (1997). Respiratory chain is required to maintain oxidized states of the DsbA-DsbB disulfide bond formation system in aerobically growing *Escherichia coli* cells. *Proc Natl Acad Sci USA* 94:11857–11862.
34. Inaba, K., Takahashi, Y.H., Ito, K., and Hayashi, S. (2006). Critical role of a thiolate-quinone charge transfer complex and its adduct form in *de novo* disulfide bond generation by DsbB. *Proc Natl Acad Sci USA* 103:287–292.
35. Bader, M., Muse, W., Ballou, D.P., Gassner, C., and Bardwell, J.C. (1999). Oxidative protein folding is driven by the electron transport system. *Cell* 98:217–227.
36. Cotter, P.A., Chepuri, V., Gennis, R.B., and Gunsalus, R.P. (1990). Cytochrome o (cyoABCDE) and d (cydAB) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the fnr gene product. *J Bacteriol* 172:6333–6338.
37. Junemann, S., Butterworth, P.J., and Wrigglesworth, J.M. (1995). A suggested mechanism for the catalytic cycle of cytochrome bd terminal oxidase based on kinetic analysis. *Biochemistry* 34:14861–14867.
38. Takahashi, Y.H., Inaba, K., and Ito, K. (2004). Characterization of the menaquinone-dependent disulfide bond formation pathway of *Escherichia coli*. *J Biol Chem* 279:47057–47065.
39. Bader, M.W., Xie, T., Yu, C.A., and Bardwell, J.C. (2000). Disulfide bonds are generated by quinone reduction. *J Biol Chem* 275:26082–26088.

40. Inaba, K., Takahashi, Y.H., Fujieda, N., Kano, K., Miyoshi, H., and Ito, K. (2004). DsbB elicits a red-shift of bound ubiquinone during the catalysis of DsbA oxidation. *J Biol Chem* 279:6761–6768.
41. Kadokura, H., Bader, M., Tian, H., Bardwell, J.C., and Beckwith, J. (2000). Roles of a conserved arginine residue of DsbB in linking protein disulfide-bond-formation pathway to the respiratory chain of *Escherichia coli*. *Proc Natl Acad Sci USA* 97:10884–10889.
42. Rietsch, A., Belin, D., Martin, N., and Beckwith, J. (1996). An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*. *Proc Natl Acad Sci USA* 93:13048–13053.
43. Sone, M., Akiyama, Y., and Ito, K. (1997). Differential *in vivo* roles played by DsbA and DsbC in the formation of protein disulfide bonds. *J Biol Chem* 272:10349–10352.
44. Darby, N.J., Raina, S., and Creighton, T.E. (1998). Contributions of substrate binding to the catalytic activity of DsbC. *Biochemistry* 37:783–791.
45. Maskos, K., Huber-Wunderlich, M., and Glockshuber, R. (2003). DsbA and DsbC-catalyzed oxidative folding of proteins with complex disulfide bridge patterns *in vitro* and *in vivo*. *J Mol Biol* 325:495–513.
46. Joly, J.C., Leung, W.S., and Swartz, J.R. (1998). Overexpression of *Escherichia coli* oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proc Natl Acad Sci USA* 95:2773–2777.
47. Kurokawa, Y., Yanagi, H., and Yura, T. (2001). Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in *Escherichia coli*. *J Biol Chem* 276:14393–14399.
48. Hoshino, K., Eda, A., Kurokawa, Y., and Shimizu, N. (2002). Production of brain-derived neurotrophic factor in *Escherichia coli* by coexpression of Dsb proteins. *Biosci Biotechnol Biochem* 66:344–350.
49. Kurokawa, Y., Yanagi, H., and Yura, T. (2000). Overexpression of protein disulfide isomerase DsbC stabilizes multiple-disulfide-bonded recombinant protein produced and transported to the periplasm in *Escherichia coli*. *Appl Environ Microbiol* 66:3960–3965.
50. Qiu, J., Swartz, J.R., and Georgiou, G. (1998). Expression of active human tissue-type plasminogen activator in *Escherichia coli*. *Appl Environ Microbiol* 64:4891–4896.
51. Schmidt, A.M., Bloss, I., and Skerra, A. (1998). Improved folding of apo-retinol-binding protein in the periplasm of *Escherichia coli*: positive influences of dsbC coexpression and of an amino acid exchange in the vitamin A binding site. *Protein Eng* 11:601–607.
52. Hu, X., O'Dwyer, R., and Wall, J.G. (2005). Cloning, expression and characterisation of a single-chai'n Fv antibody fragment against domoic acid in *Escherichia coli*. *J Biotechnol* 120:38–45.
53. Bessette, P.H., Cotto, J.J., Gilbert, H.F., and Georgiou, G. (1999). *In vivo* and *in vitro* function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. *J Biol Chem* 274:7784–7792.
54. Shao, F., Bader, M.W., Jakob, U., and Bardwell, J.C. (2000). DsbG, a protein disulfide isomerase with chaperone activity. *J Biol Chem* 275:13349–13352.
55. Chen, J., Song, J.L., Zhang, S., Wang, Y., Cui, D.F., and Wang, C.C. (1999). Chaperone activity of DsbC. *J Biol Chem* 274:19601–19605.
56. Berkmen, M., Boyd, D., and Beckwith, J. (2005). The nonconsecutive disulfide bond of *Escherichia coli* phytase (AppA) renders it dependent on the protein-disulfide isomerase, DsbC. *J Biol Chem* 280:11387–11394.
57. Hiniker, A., and Bardwell, J.C. (2004). *In vivo* substrate specificity of periplasmic disulfide oxidoreductases. *J Biol Chem* 279:12967–12973.
58. Wunderlich, M., and Glockshuber, R. (1993). Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci* 2:717–726.

59. Zapun, A., Missiakas, D., Raina, S., and Creighton, T.E. (1995). Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry* 34:5075–5089.
60. Heras, B., Edeling, M.A., Schirra, H.J., Raina, S., and Martin, J.L. (2004). Crystal structures of the DsbG disulfide isomerase reveal an unstable disulfide. *Proc Natl Acad Sci USA* 101:8876–8881.
61. McCarthy, A.A., Haebel, P.W., Torronen, A., Rybin, V., Baker, E.N., and Metcalf, P. (2000). Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat Struct Biol* 7:196–199.
62. Tian, G., Xiang, S., Noiva, R., Lennarz, W.J., and Schindelin, H. (2006). The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* 124:61–73.
63. Segatori, L., Paukstelis, P.J., Gilbert, H.F., and Georgiou, G. (2004). Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in *Escherichia coli*: Rreconciling two competing pathways. *Proc Natl Acad Sci USA* 101:10018–10023.
64. Bader, M.W., Hiniker, A., Regeimbal, J., Goldstone, D., Haebel, P.W., Riemer, J., Metcalf, P., and Bardwell, J.C. (2001). Turning a disulfide isomerase into an oxidase: DsbC mutants that imitate DsbA. *EMBO J* 20:1555–1562.
65. Rietsch, A., Bessette, P., Georgiou, G., and Beckwith, J. (1997). Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J Bacteriol* 179:6602–6608.
66. Stewart, E.J., Katzen, F., and Beckwith, J. (1999). Six conserved cysteines of the membrane protein DsbD are required for the transfer of electrons from the cytoplasm to the periplasm of *Escherichia coli*. *EMBO J* 18:5963–5971.
67. Chung, J., Chen, T., and Missiakas, D. (2000). Transfer of electrons across the cytoplasmic membrane by DsbD, a membrane protein involved in thiol-disulphide exchange and protein folding in the bacterial periplasm. *Mol Microbiol* 35:1099–1109.
68. Katzen, F., and Beckwith, J. (2000). Transmembrane electron transfer by the membrane protein DsbD occurs via a disulfide bond cascade. *Cell* 103:769–779.
69. Collet, J.F., Riemer, J., Bader, M.W., and Bardwell, J.C. (2002). Reconstitution of a disulfide isomerization system. *J Biol Chem* 277:26886–26892.
70. Haebel, P.W., Goldstone, D., Katzen, F., Beckwith, J., and Metcalf, P. (2002). The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: Crystal structure of the complex. *EMBO J* 21:4774–4784.
71. Rozhkova, A., Stirnimann, C.U., Frei, P., Grauschopf, U., Brunisholz, R., Grutter, M.G., Capitani, G., and Glockshuber, R. (2004). Structural basis and kinetics of inter- and intramolecular disulfide exchange in the redox catalyst DsbD. *EMBO J* 23:1709–1719.
72. Katzen, F., and Beckwith, J. (2003). Role and location of the unusual redox-active cysteines in the hydrophobic domain of the transmembrane electron transporter DsbD. *Proc Natl Acad Sci USA* 100:10471–10476.
73. Pan, J.L., and Bardwell, J.C.A. (2006). The origami of thioredoxin-like folds. *Protein Sci* 15:2217–2227.

# 6

# *The Identification of the YaeT Complex and Its Role in the Assembly of Bacterial Outer Membrane β-Barrel Proteins*

JULIANA MALINVERNI • THOMAS J. SILHAVY

*Department of Molecular Biology*
*Princeton University*
*Princeton, NJ 08544, USA*

## I. Abstract

The outer membrane (OM) of Gram-negative bacteria is an essential organelle that prevents the indiscriminate diffusion of small molecules while permitting the selective entry of vital nutrients. To achieve this barrier function, the lipid to protein ratio of the OM must be carefully maintained despite dynamic growth conditions. Severe alterations in envelope composition lead to loss of OM integrity and cell death, a fact that had for decades hampered the search for viable mutations in genes responsible for the expression of OM assembly factors. Moderate changes in composition lead to increased OM permeability, and this general characteristic can result from an array of different mutations.

Genetic characterization of hyperpermeable mutants can be complicated because the resulting phenotypes are often subtle and pleiotropic. Recent evidence demonstrates that specific mutations that alter OM permeability become distinguishable through the use of toxic small molecules.

THE ENZYMES, Vol. XXV

ISSN NO: 1874-6047

DOI: 10.1016/S1874-6047(07)25006-1

This exciting realization allowed the design of powerful selections to probe for genes expressing factors involved in OM biogenesis. Bacteria carrying undefined OM permeability mutations can be used in selections that demand a distinct and reproducible set of genetic solutions that restore OM barrier function and prevent specific toxic compounds from reaching an intracellular target. We review how this discovery led to the identification of YaeT (Omp85), an ancient OMP assembly factor conserved in Gram-negative bacteria and eukaryotes. Moreover, we evaluate the genetic and biochemical interactions between YaeT and other components that also play a role in the translocation and assembly of OMPs in the OM.

## II. Gram-Negative Bacterial Envelope

The envelope of Gram-negative bacteria provides structural integrity to the cell, regulates the flux of solutes, and serves as the site for a number of enzymes involved in metabolic processes such as DNA replication and cell division, signal transduction, electron transport, oxidative phosphorylation, and the synthesis of the cell wall itself. The envelope consists of three distinct compartments: an inner membrane (IM), an aqueous periplasm containing a thin peptidoglycan layer, and an outer membrane (OM). The IM is a typical symmetric bilayer of phospholipids, while the OM is an asymmetric bilayer with an inner phospholipid leaflet facing the periplasmic compartment and an outer leaflet composed primarily of lipopolysaccharides (LPS) [1]. In addition to the lipid and peptidoglycan components, the bacterial envelope also contains multiple classes of proteins, which can be either peripherally or integrally associated with the membranes or freely soluble in the periplasm [1].

Precursor envelope components are synthesized in the cytosol and subsequently undergo modifications at the inner or outer leaflet of the IM. Once translocated across the cytoplasmic membrane, amphipathic lipids and proteins destined for the outermost bilayer must transcend several daunting challenges prior to reaching their final destinations. These molecules must bypass the rigid peptidoglycan barricade, traverse the aqueous periplasmic space, and assemble into the OM all in an environment lacking an obvious energy source such as ATP.

## III. Protein Transport Across the Bacterial Envelope

The OM contains two classes of proteins: the integral OM β-barrel proteins (OMPs) that span the bilayer, and OM lipoproteins that are tethered to the inner leaflet of the OM by covalently attached lipid, but reside mainly within the periplasmic space. Both types of proteins are

synthesized in the cytoplasm and directed to the IM SecYEG translocation complex via their N-terminal signal sequences (see [2], for review) Once translocation from the cytoplasm is completed, the fates of the two classes of proteins diverge. The processing and transit of lipoproteins to the OM has been well characterized in *Escherichia coli*, and will be described in more detail elsewhere in this volume (also, see [3] and references therein). Briefly, a precursor lipoprotein is modified at the periplasmic face of the IM subsequent to Sec translocation, where a requisite cysteine residue adjacent to the signal sequence is converted to a thioetherdiglyceride. The lipoprotein is then cleaved of the signal sequence by a dedicated signal peptidase and amino-acylated at the modified cysteine residue now located at the newly exposed amino terminus. IM lipoproteins contain an aspartate residue at the +2 position of the processed molecule that allows these proteins to remain tethered at the periplasmic face of the IM via the N-terminal fatty acid moieties. An OM lipoprotein lacks the aspartate at the +2 position, and as a result is recognized by the ATP-binding cassette (ABC) transporter, LolCDE. The LolCDE complex then releases the lipoprotein in an ATP-dependent manner to the periplasmic chaperone LolA that keeps the hydrophobic moieties in a soluble state during periplasmic transit. Finally, LolA passes the OM lipoprotein onto the essential OM lipoprotein LolB, which acts as a receptor that presumably guides the attachment of lipoproteins to the phospholipid leaflet of the OM [3].

Determination of the OMP assembly pathway has remained more elusive than the lipoprotein assembly pathway. Once OMPs and periplasmic proteins are translocated across the cytoplasmic membrane through the SecYEG translocation complex their signal sequences are cleaved [2], effectively releasing periplasmic proteins into the hydrophilic matrix. The fate of processed OMPs is less clear. Two prevailing models have been proposed to explain how OMPs are routed across the periplasmic space. The ''periplasmic intermediate'' model proposes that OMPs are targeted to the OM across the periplasm via a soluble OMP-chaperone complex (Figure 6.1A). The ''Bayer's patches'' model contends that OMP transit occurs via putative zones of adhesion between the IM and OM without the presence of soluble periplasmic intermediates (Figure 6.1B). Evidence exists for both models, and neither can be easily discounted.

One of the strongest arguments for the periplasmic intermediate model is that soluble OMPs can be isolated in the periplasmic space *in vivo*. However, these findings were mostly drawn from experiments in which the OMPs in question were overproduced, and periplasmic localization under these conditions may be a reflection of a saturated primary pathway rather than a preferred assembly pathway [4, 5]. To our knowledge no protein has been shown to reliably ''chase'' from an unfolded form in the periplasm into a stably folded OMP in the OM with the exception of TolC,

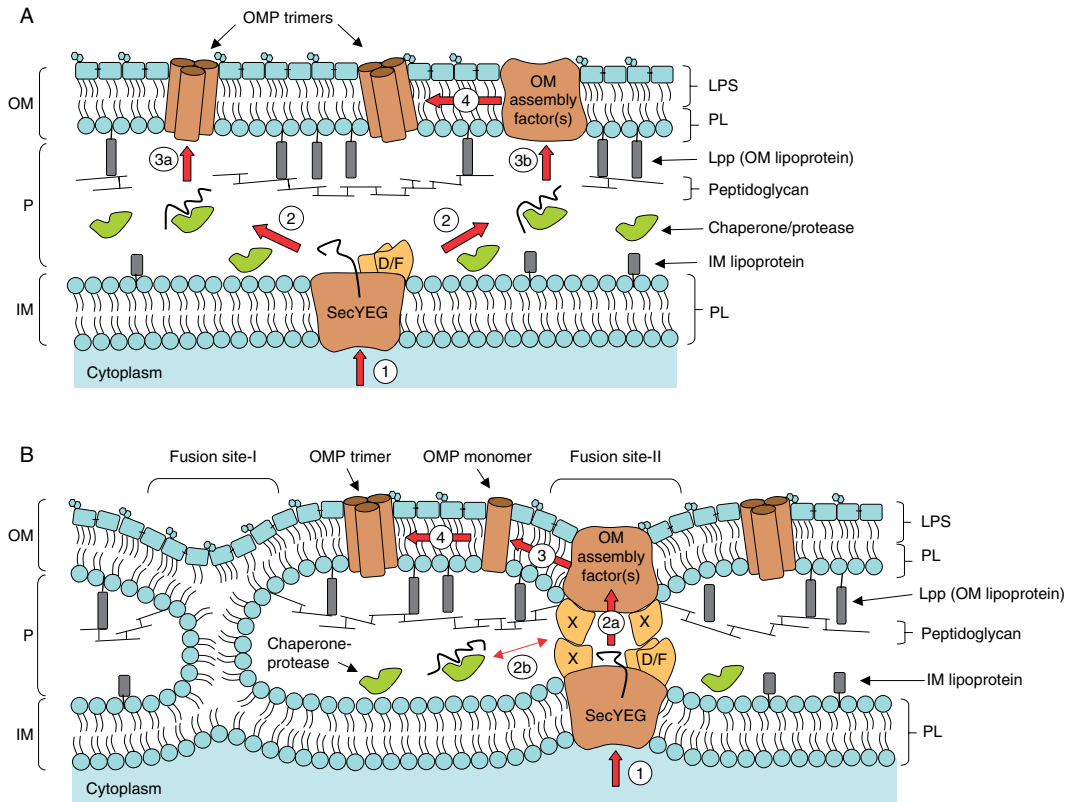



FIG. 6.1. Transport models across and into bacterial membranes. (A) Periplasmic intermediate model. Translocation of the unfolded OMP intermediate (shown as a bold black line) occurs through the SecYEG translocon (step 1). SecD and SecF (labeled D/F in the sketch) appear to facilitate translocation and/or release of OMPs from the IM [2]. Subsequent to the processing of the signal sequence, the unfolded mature protein

a unique extended α-β-barrel trimeric porin [6]. However, it may be that the transit and folding of most conventional transmembrane β-barrels *in vivo* is too rapid a process to isolate any true periplasmic OMP intermediates under normal expression conditions. Another central argument for the periplasmic intermediate model comes from studies that have demonstrated that nascent OMPs can be excreted into the medium from spheroplasts [7, 8]. This model proposes that once released into the periplasm, OMPs bind to periplasmic chaperones and foldases which solubilize OMPs and direct their transit across the periplasm (Figure 6.1A) (see [9], for review). Indeed, extracellular stresses can result in the accumulation of misfolded OMPs in the periplasmic space, which in turn requires the cell to upregulate the production of periplasmic chaperones and proteases to prevent cell death [9, 10]. However, whether periplasmic OMPs are merely dead-end molecules that escape into the matrix on excess production or stressful conditions, or whether they are indicative of true assembly intermediates temporarily halted during periplasmic transit remains unclear.

The Bayer's patches model evokes transport across zones of adhesion between the OM and IM, most likely proteinaceous in nature, and stems

binds to a periplasmic chaperone such as Skp or SurA, which ferries the intermediate from the IM across the periplasm (step 2) [9]. Some variations of the model assert that OMP assembly into the OM occurs spontaneously, and the reaction made favorable with the help of LPS, phospholipid, and/or interactions with the periplasmic chaperone foldases (step 3a). Alternatively, the periplasmic chaperones deliver the OMP substrate to an OM assembly site (step 3b). This assembly machine may either indirectly catalyze the OMP insertion reaction by providing localized chaperone activity near the inner leaflet of the OM, or it may constitute a translocon analogous to the IM Sec/SRP pathway that requires transport of OMP substrate through a porin chamber followed by lateral release of folded mature protein (steps 3b and 4) [2]. (B) Bayer's Patch Model. This model evokes putative contact sites between the IM and OM [11]. Although these sites may be lipid in nature (as shown in fusion site – I), we favor the hypothesis that contact sites would most likely be proteinaceous in nature (as shown in fusion site – II). In this model, the Sec translocon transiently docks to the OM assembly machinery. Transport of the OMP precursor across the IM occurs as described above and in the text (step 1). Mature unfolded OMP substrates then cross the periplasmic barrier through a proteinaceous channel composed of unknown proteins (depicted as "X" in the sketch) (step 2a). At this point, it is conceivable that OMPs could escape the "transport tunnel" into the periplasmic space on saturation of the assembly pathway, where they are solubilized by periplasmic chaperones (step 2b). Successful translocation events across the protein channel would result in the recognition of the unfolded OMP by OM assembly factor(s) and subsequent assembly into the OM (step 3). The process by which this may occur is unknown (see text). (A and B) Many proteins exist in a trimeric state in the OM, and it is unknown how this association is initiated. It may be that trimerization of monomeric subunits is catalyzed by a series of monomer–lipid interactions, or it may be that an additional OM assembly factor is required for multimerization (trimerization depicted in panel B, step 4).

from observations by Manfred Bayer of localized patches in the cell wall of *E. coli* visualized by electron microscopy (Figure 6.1B) [11]. Indeed, *de novo* assembly of LPS and porins has been observed to occur at regions that colocalize with such patches in *Salmonella typhimurium* and in *E. coli*, respectively [12, 13]. Another study used fractionation techniques to demonstrate that radiolabeled substrates pass directly from the cytoplasmic fraction to the OM fraction without delay [14]. However, this transit model is not without its critics. The very existence of Bayer's patches was called into question when patches could not be observed using alternative methods of microscopy sample preparation [15]. However, the theory of fusion sites between membranes was recently bolstered by the finding that newly synthesized LPS molecules localize to the OM in spheroplasts. Furthermore, the LPS could not be released from the spheroplasts by treatment with periplasmic fractions [16]. In other words, the transport of LPS does not appear to require a periplasmic chaperone for release into the periplasm, rather localization of LPS proceeds in the absence of an obvious soluble intermediate. Evidence for such a model is not without precedent, since protein import across the double bilayer of eukaryotic mitochondria has been shown to proceed via transient contacts between protein machineries of the IM and OM of these organelles formed during translocation (see [17], for review).

Ultimately, we do not yet know if either of the models accurately represents the primary pathway that OMPs undergo to travel across the bacterial periplasm. One could envision a combination of the models proposed. For example, OMPs could follow a translocation route via protein channels connecting the two membranes, but reversibly associate with periplasmic factors on abortive assembly attempts or ''backing up'' of the transit system. Furthermore, any model must take into account transit across the peptidoglycan, which could conceivably occur either through gaps in the peptidoglycan scaffolding or a bypass of the layer via proteinaceous adhesion zones (Figure 6.1A and B) [18]. In any event, OMP substrates must undergo assembly and insertion once they reach the OM. Until recently, it was not known if this process occurred spontaneously or if a dedicated assembly factor(s) was required for OMP assembly (Figure 6.1A). Indeed, some OMPs have been shown to spontaneously fold in the presence of some combination of LPS, periplasmic factors, and/or phospholipids *in vitro* [19–21]. Although evidence suggests that these factors may indeed play legitimate roles in OMP assembly, it was noted in some of these same studies that the kinetics of spontaneous assembly are too slow to account for *in vivo* observations and other factors were likely to be involved.

## IV. Identification of OM Biogenesis Factors: The Search for Needles in a Haystack

The quest for OM biogenesis factors using conventional genetic techniques has historically proven to be an elusive challenge in *E. coli* due to the essential nature of the envelope and the lack of appropriate selections. To narrow the search for potential candidates, our laboratory sought to establish a set of criteria that we predicted would hold true for all OM assembly factors. The profile delineated four anticipated hallmarks. The first and most obvious stipulation was that any major assembly factor would be associated with the OM and it would have to be essential, since the OM is required for viability. Second, should a partial loss-of-function mutation be discovered in a gene encoding a putative OM assembly factor, it would be expected to alter barrier function and cause measurable changes in OM permeability. Third, we expected that a gene encoding an OM assembly factor would be regulated by the $\sigma^E$ stress response, which is essential for growth [22] and is involved in the maintenance of homeostatic control of the envelope under a variety of conditions [23]. Finally, we postulated that a gene essential for the assembly of the OM would likely be conserved across most, if not all, Gram-negative genera, and the gene would tend to cluster near other genes expressing proteins of related function.

According to our profile approach, one gene in particular seemed to fit the above criteria closely enough to warrant further investigation. This gene, designated for *i*ncreased *m*embrane *p*ermeability (*imp*), was first characterized in *E. coli* in 1989 in a genetic selection designed to search for mutations that would increase OM permeability enough to allow large maltodextrin molecules to freely diffuse across the barrier [24]. Mutations in the *imp* locus that conferred this hyperpermeable phenotype were clearly not null mutations, and the authors of this work postulated that the gene might be essential [24]. In addition, it had been demonstrated that the *imp* gene was expressed from a $\sigma^E$-dependent promoter [25]. Moreover, the *imp* locus lay directly upstream of *surA*, a periplasmic chaperone with peptidyl-prolyl isomerase activity that had been demonstrated to aid in the assembly of some OMPs [26, 27].

Further examination of the *imp* gene by our laboratory confirmed that it is indeed an essential gene in *E. coli*, and that the protein encoded by *imp* is a large OM β-barrel protein [28]. The *imp* gene was also found to be conserved in many Gram-negative bacteria [28]. Evidence that Imp is involved in envelope biogenesis came from depletion studies; cells depleted of Imp produce an abnormally high density membrane fraction consistent with an envelope assembly defect [28].

Other studies published since have strengthened the validity of the profile approach in the search for an OM biogenesis factor. It was demonstrated that Imp is involved in the transport of LPS to the cell surface in *Neisseria meningitidis* [29]. Furthermore, $\sigma^E$-regulation of *imp* has since been confirmed by other groups [30, 31], the importance of which is clarified by the finding that the stress response is not only induced in response to misfolded OMPs, but to alterations in LPS as well [32]. It appears that the essential role of $\sigma^E$ is not limited to monitoring perturbations in the OM, but that it may in fact coordinate a transcriptional network required for maintaining the availability of envelope precursors throughout growth [31].

Using a biochemical approach, our laboratory in collaboration with the Kahne laboratory identified an OM lipoprotein, annoted as *r*are *l*ipo*p*rotein *B* (RlpB) [33], that copurifies with Imp [34]. We found that RlpB matched at least two criteria outlined in our profile approach; the gene is essential; and it is conserved in most Gram-negative bacteria, most often occurring in organisms that also contain Imp ([34], unpublished observations). It remains to be determined if the *rlpB* locus is regulated by $\sigma^E$, but it is possible the gene has avoided detection in screens for $\sigma^E$-regulated promoters as the low levels of protein in the cell may indicate low levels of transcription at this promoter, even when induced. In addition, we do not yet have a partial loss-of-function mutation in *rlpB*, but we would predict that such a mutation would cause defects in OM permeability similar to those conferred by *imp* partial loss-of-function mutations.

Depletion of RlpB results in phenotypes similar to those seen on Imp depletion; the accumulation of mistargeted LPS and the consequent incorporation of phospholipids into the outer leaflet of the OM, the production of ''extra'' membranes, and ultimately cell death [28, 29, 34]. The precise role of RlpB is unknown; however, we know it is not required for the stabilization of Imp, and that depletion of either Imp or RlpB has little affect on OMP assembly [29, 34]. RlpB contributes in a significant way to the LPS assembly pathway, and together with Imp functions to ensure proper LPS localization [34].

In Section V, we describe a chemical genetic technique for identifying additional genes for OM biogenesis factors. Many of the genes identified closely fit the profile described above demonstrating the utility of these criteria.

## V. Chemical Conditionality: The YfgL Connection to OM Assembly

As described earlier, the *imp* locus was first discovered using a selection to find mutants with increased OM permeability. Strains carrying one such allele, *imp4213*, are hypersensitive to a wide array of toxic small molecules,

including bile salts, vancomycin, and a hydrophobic derivative of vancomycin, chlorobiphenyl-vancomycin (CBPV) [24, 35]. Our laboratory in collaboration with the Kahne laboratory sought to find suppressors of *imp4213* by selecting for resistance to these compounds. We found that the quality and quantity of suppressor types varied depending on the selection pressure that was applied [36]. A high quality suppressor is defined as one that restores the OM permeability barrier to near wild-type effectiveness, whereas a low quality suppressor is one that improves the barrier enough to prevent entry of only a small subset of toxic molecules. The quantity of suppressor types refers to the number of mutation classes that answered a given selection.

Vancomycin elicited highly specific selection pressure for *imp4213* suppressors because only one type of suppressor answered the selection for vancomycin resistance, intragenic mutations in *imp4213* itself. The quality of these suppressors was high, as they could also confer a level of resistance to all of our test molecules. Suppressors that arose on CBPV selection on the other hand yielded only one additional type of suppressor, extragenic mutations that lowered or abolished the production of YfgL, an OM lipoprotein of unknown function. These suppressors did not restore barrier function sufficiently enough to allow growth in the presence of vancomycin, but surprisingly, they did restore the ability to grow in the presence of bile salts. This important observation showed that YfgL was not itself a target of the CBPV, but that disruption of YfgL improved the overall quality of the *imp4213* OM in such a way as to prevent the structurally different CBPV and bile salts molecules from crossing the barrier, but not enough to block vancomycin entry. Finally, bile salts applied the least specific selection pressure of our test compounds, and resulted in multiple classes of *imp4213* suppressors. About half of the suppressor mutations were in *yfgL*, and could therefore also confer CBPV resistance, but the remaining suppressors could confer resistance to bile salts but not to CPBV or vancomycin. This ''chemical conditionality'' approach elicited the continuum of *imp4213* suppressor classes depicted schematically in Figure 6.2 [36].

Strikingly, loss-of-function mutations in *yfgL* in cells that are wild type for the *imp* allele compromise barrier function and also confer sensitivity to CBPV, but not to bile salts. Thus both *yfgL* and *imp4213* impair barrier function but the mutations together improve barrier function. In a sense, each mutation is a suppressor of the other. This suggests that YfgL and Imp participate in different but related biochemical processes [36]. The genetic connection between *yfgL* and *imp* and the fact that mutations in *yfgL* affect OM permeability strongly implied that YfgL might also be involved in OM biogenesis. Indeed, an additional phenotype of *yfgL* mutants is decreased steady state levels of OMPs [37]. Furthermore, it has since been shown that *yfgL* is regulated by the $\sigma^E$ pathway [30, 31, 38].

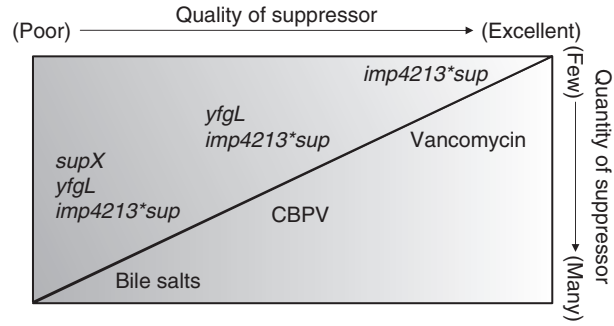



FIG. 6.2. Continuum of suppressor quality and quantity elicited by chemical conditionality. The set of chemicals elicited in our chemical conditionality approach were bile salts, CBPV, and vancomycin. Exposure of a hypersensitive mutant, *imp4213*, to these compounds allowed us to select for suppressors to each of these compounds. These suppressors do not alter the mode of action of the compounds at their intracellular targets rather they alter the ability of these compounds to reach their targets by improving the quality of the envelope to various degrees. Vancomycin selection results in only intragenic *imp4213* suppressors (*imp4213*sup*), and these mutations suppress sensitivity to CBPV and bile salts. CBPV elicits one additional suppressor, *yfgL*, which is of intermediate quality and allows *imp4213* to grow in the presence of CBPV and bile salts, but not vancomycin. Bile salts selection yields the highest number of suppressors, approximately half of which are in *yfgL*, while the other half are composed of a mixture of suppressors of relatively poor quality (*supX*) that only confer resistance to bile salts and not CBPV or vancomycin [36].

YfgL meets some of the criteria highlighted in our profile approach for OM assembly factors. However, *yfgL* is not essential and the locus does not appear to cluster with other known OM biosynthesis genes. Likely reflecting the fact that YfgL is a nonessential OM assembly factor, we find that there are homologues of *yfgL* scattered among OM-containing bacteria, but a recognizable *yfgL* homologue is absent in several representative groups of Gram-negative bacteria and the gene does not necessarily co-occur in genomes that contain *imp* (unpublished observations).

## VI. Identification and Characterization of the YaeT Complex

The genetic link between *imp* and *yfgL* implied a biochemical connection that may exist between the encoded proteins, and we set out to determine if His-tagged YfgL would copurify with Imp [37]. Remarkably, we could not detect an interaction between this OM lipoprotein and Imp, but instead found that YfgL interacts with an OMP, YaeT, and two other OM lipoproteins, YfiO and NlpB. This was an exciting discovery on

multiple levels. Most intriguing was the YaeT protein because published data available at the time suggested that this protein fit at least three of the four criteria that we predicted would describe an OM assembly factor. First, homologues of *yaeT* were hypothesized or proven to be essential in other organisms [39–42], although conflicting results clouded the essential nature of *yaeT* in *E. coli* [25, 43]. Second, *yaeT* was reported to be regulated by $\sigma^E$ [25] and was subsequently confirmed by a later study [31]. Third, the *yaeT* locus lies directly upstream of *skp*, a periplasmic chaperone [9]; furthermore, both genes are surrounded by a cluster of LPS biosynthesis genes, a chromosomal arrangement that is preserved in multiple bacterial genera [42]. Moreover, homologues of *yaeT* are found in virtually every Gram-negative bacterium OM [42, 44] (except for perhaps the primitive OMs of Chlorobacteria, or ''green non-sulfur'' bacteria [45]), and *yaeT* homologues are present in eukaryotic mitochondria and chloroplasts [44].

The Wagner and Keegstra laboratories first noted that a cyanobacterial homologue of *yaeT*, SynToc75, shared structural and possibly functional homology with a protein importer located in the OM of chloroplasts, Toc75 [39, 46]. This finding was later extended by the van der Ley and Tommassen groups to studies of a YaeT homologue in *N. meningitidis*, designated Omp85 [41, 42]. The two groups independently demonstrated that depletion of Omp85 led to severe alterations of the neisserial envelope; however, they reached conflicting conclusions: the van der Ley group concluded that depletion of Omp85 interfered with LPS assembly [41], whereas the Tommassen group concluded that depletion of Omp85 interfered with OMP assembly [42]. It turns out that the Tommassen group was correct. In the particular strain of *Neisseria* in question, LPS is not essential. Imp is also not essential in this strain [47], but Omp85 is [41, 42]. This reflects the fact that Imp is required for LPS assembly while Omp85 is not [42]. The Tommassen laboratory demonstrated that depletion of Omp85 leads to an accumulation of misassembled OMPs, and that Omp85 can directly bind to multiple OMP substrates [42]. Moreover, the YaeT/Omp85 homologues that are conserved in the OM of mitochondria, termed Sam50 or Tob55 in *Saccharomyces cerevisiae* and *Neurospora crassa*, respectively, have also been shown to be responsible for the assembly of β-barrel proteins in the OM of these eukaryotic organelles (see [48], and references therein).

In the meantime, our laboratory in conjunction with the Kahne laboratory set out to determine what effect, if any, the depletion of YaeT would have on the OM of *E. coli* [37]. Our results confirmed that *yaeT* is indeed essential in *E. coli.* We demonstrated that depletion of YaeT led to a drastic reduction in the overall steady state levels of OMPs. Moreover, over the course of time required for YaeT depletion, the levels of folded OMPs decreased and this correlated with an increase in the level of misfolded

OMPs [37]. Other groups have since confirmed our results, showing that YaeT affects the assembly of all OMPs tested to date in *E. coli*, and has little to no influence on the assembly of LPS [49, 50].

Incidentally, we have recent evidence that allows YaeT to fulfill the last remaining prediction of our profile approach for an essential OM biogenesis factor, that partial loss-of-function mutations in such a gene would lead to OM permeability defects. Characterization of one of the relatively ''poor quality'' suppressors of *imp4213* in the presence of bile salts (section V) led to the discovery of a partial loss-of-function mutation in YaeT, designated YaeT6 [51]. Although this mutation improves the quality of the *imp4213* envelope, the *yaeT6* allele in an *imp*$^+$ background demonstrates increased OM sensitivity compared to the wild-type envelope [51]. The *yaeT6* allele exerts its phenotype independently of *yfgL*, as it has been shown that YaeT6 does not render YfgL nonfunctional [51].

What effect do the OM lipoproteins associated with YaeT have on the integrity of the OM, and what correlations can we make with our OM assembly factor profile? All three lipoproteins are regulated by the $\sigma^E$ pathway in *E. coli* [25, 30, 31, 38], but only YfiO is essential for viability [38, 52]. Strains producing C-terminal truncations of YfiO are still viable and exhibit reduced levels of OMPs and an increase in OM permeability similar to that of *yfgL* mutants [37]. Depletion of YfiO leads to similar OMP defects as YaeT depletion, and YfiO depletion also has negligible effects on LPS assembly [52]. This argues that YfiO and YaeT are integral components of the OMP assembly machine in *E. coli* [52]. The importance of YfiO is underscored by its high degree of conservation in most sequenced representatives of Gram-negative bacteria [52]; however, there are some examples of Gram-negative phyla that do not appear to contain an obvious homologue of YfiO (unpublished observations).

NlpB, on the other hand, is the least ubiquitous of the YaeT complex members among Gram-negative bacteria. Homologues of NlpB do not appear to be present outside of the $\gamma$ and $\beta$ subphyla of Proteobacteria (unpublished observations). Phenotypes of *nlpB* mutants are extremely subtle under normal laboratory conditions; reduction of OMPs and OM permeability defects are reproducibly minor compared with wild type [37]. The severity of *nlpB* defects becomes clear only when combined with other YaeT-complex mutants. For example, *yfgL*$^-$ *nlpB*$^-$ double mutants exhibit irregular colony morphology and grow more slowly than either of the individual mutants ([37], unpublished observations). Therefore, all three OM lipoproteins of the YaeT complex appear to play some role in OMP assembly, although the individual contributions of each factor are not yet known. YfgL and NlpB appear to be auxiliary OM assembly factors, while

YfiO, like YaeT, is required absolutely for OM biogenesis in *E. coli*, and likely many other Gram-negative bacteria as well.

## VII. Interactions Among YaeT Complex Members

Biochemical immunoprecipitation of any member of the YaeT complex results in the copurification of all of the remaining components of the complex [37]. Using mutations that alter the lipoprotein components of the complex, we could further probe which interactions are direct and which are indirect [52]. For example, purification of the YaeT complex in a $YfgL^-$ background has no effect on the ability of the rest of the complex components to copurify. Similarly, purification of the YaeT complex in an $NlpB^-$ context has no effect on the ability of the remaining members of the complex to copurify. However, truncations in the C-terminus of YfiO ($YfiO_{\Delta C\text{-term}}$) result in weakened interactions between NlpB and the rest of the YaeT complex, but do not affect the ability of YfgL to bind YaeT. Therefore, it appears that YfgL and YfiO each directly interact with YaeT at distinct sites, and that interaction of NlpB with YaeT may be indirect and requires full-length YfiO protein to produce stable contacts with the complex (Figure 6.3) [52].

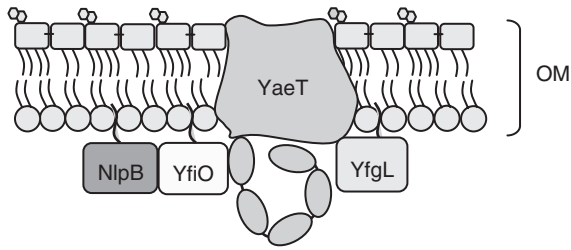



FIG. 6.3. Proposed physical organization of the YaeT complex. Biochemical coprecipitation experiments indicate that YfiO and YfgL each directly contact independent sites of the YaeT protein. The C-terminus of YfiO is required for strong interactions between NlpB and the rest of the YaeT complex, indicating that NlpB and YfiO directly interact. Genetic evidence demonstrates that NlpB must share at least one additional contact with the YaeT complex other than with the C-terminus of YfiO. While the identity of this additional contact remains unknown, it is unlikely to be with YfgL (see text) [52]. The five POTRA domains of YaeT are indicated in ovals, and the C-terminal β-barrel domain is depicted as residing within the OM. The sketch is meant to demonstrate the existence of biochemical interactions among the OM lipoproteins and YaeT, and not the precise domains of YaeT with which YfiO and YfgL interact.

One explanation for the above observations is that the C-terminus of the YfiO protein is the sole point of contact required for NlpB interactions with the complex. If this were true, then assuming that NlpB requires proper localization to the YaeT complex for function, we would predict that the $YfiO_{\Delta C\text{-term}}$ mutant would exhibit the same phenotype as a $YfiO_{\Delta C\text{-term}}$ $NlpB^-$ double mutant. However, the $YfiO_{\Delta C\text{-term}}$ $NlpB^-$ double mutant is temperature sensitive whereas either single mutant is not. In addition, the $YfiO_{\Delta C\text{-term}}$ $NlpB^-$ double mutant has lowered levels of OMPs and increased OM permeability relative to the $YfiO_{\Delta C\text{-term}}$ $NlpB^+$ mutant ([52], unpublished observations). Therefore, we presume that the C-terminus of YfiO is necessary for a strong interaction with NlpB, but that one or more additional contacts may exist between NlpB and the YaeT complex. These additional contacts may include interactions with YaeT, some other as yet uncharacterized component of the complex, or some portion of YfiO other than the C-terminus. We do not believe the additional contact(s) are with YfgL, given that loss of the YfgL lipoprotein has no affect on the YfiO/NlpB subset of interactions with YaeT.

## VIII. POTRA Domains

As mentioned previously, YaeT belongs to a larger family of proteins that appear to share a role in OMP assembly or transport. This evolutionarily conserved group, referred to as the ''bacterial-surface antigen'' superfamily, can be separated into two main categories [48], those that assemble OMPs into the OM of a cell or organelle, and those that transport peptide substrates across the OM. The former group includes all YaeT/Omp85 bacterial homologues as well as the Tob55/Sam50-like proteins found in the OMs of all plant and animal mitochondria [48]. The second group of bacterial-surface antigen proteins is represented by two distinct OMP families; the Toc75 proteins which form a translocation channel as part of a general protein import pathway into chloroplasts [39, 46], and bacterial OMPs that belong to the two-partner secretion systems (TPS), also known as the type V secretion pathway [48, 53, 54]. The OMP components of the TPS pathway (generally referred to as TpsB proteins) are inserted into the OM and subsequently secrete a large exoprotein (generally referred to as TpsA proteins) across the OM [53]. Although the TpsB components of the TPS pathway share structural and possible functional similarities to the more distantly related YaeT/Omp85 bacterial homologues, they are not present in most Gram-negative bacteria and are not essential in *E. coli* [48, 53].

What all members of the bacterial-surface antigen superfamily share in common is a C-terminal OM β-barrel domain and a large N-terminal periplasmic domain containing variable numbers of a structurally distinct feature designated a *PO*lypeptide *TR*anslocation *A*ssociated (POTRA) domain (Figure 6.3) [42, 55]. POTRA domains are predicted to consist of three β-strands with two α-helices occurring between the second and third β-strands (β−β−α−α−β) [55]. For example, YaeT and Omp85 are predicted to contain five POTRA domains each, while Sam50/Tob55 and TpsB proteins contain only one such domain [48, 55]. Interestingly, the only other group of proteins aside from the bacterial-surface antigen superfamily that has been predicted to carry a POTRA domain is the FtsQ/DivIB family of cell division proteins [55]. These proteins are predicted to contain a single periplasmic POTRA domain, but unlike members of the bacterial-surface antigen superfamily, are tethered to the IM through a single helical transmembrane domain [56].

At the time of this writing, no structure containing a POTRA fold has been published, but several observations have been made that may give us some insight into the possible functions of POTRA domains. Available evidence indicates that these regions are involved in the direct binding of substrates. DivIB has been suggested to act as a chaperone to stabilize the intrinsically unstable divisomal protein FtsL in *Bacillus subtilis* [56], and it has been proposed that the POTRA domain of DivIB is specifically responsible for this activity [57]. Furthermore, Toc75 of *Pisum savitum* plastids were shown to require the N-terminal POTRA domain for substrate binding, whereas a more extreme N-terminal region outside of the POTRA domain appeared to be required for complex formation of the different components of the TOC translocon machinery [58].

We know that YaeT-like proteins also physically interact with purified OMP substrates [42, 59]. If the POTRA domains of YaeT-like proteins are responsible for the binding activity, it would be interesting to determine if the interactions are required for OMP folding and/or reorganization of OMP precursors prior to OM insertion, or whether the POTRA domains merely provide a chaperone function in the form of a soluble scaffold prior to the folding and assembly of OMP intermediates by the β-barrel portion of YaeT. Although the function of the POTRA domains is unknown, the physiological relevance of these domains is evidenced by the fact that their truncation from the β-barrel portion of YaeT leads to loss of viability in *E. coli* [60]. It is also interesting to note that the *yaeT6* partial loss-of-function mutation (section VII) is an in-frame insertion of two amino acids between codons 218 and 219 of YaeT, and is located just upstream of the first of the two predicted α-helices within POTRA domain number 3 [51].

The YaeT6 mutation leads to lower levels of OMPs and suggests that alteration of a single POTRA domain can impair YaeT function [51].

## IX. Properties of the YaeT-Like β-Barrel Domains

The β-barrel domain of YaeT is predicted based on sequence analysis of Omp85 to consist of 12 transmembrane strands [42]. It is possible that the C-terminal β-barrel portion of the YaeT-like proteins serves no additional function other than to anchor the periplasmic business end of the protein in the correct orientation and provide functional interactions with other components of the YaeT complex. Alternatively, the β-barrel domain could promote a critical step in the OMP assembly/translocation pathway, either separately or in conjunction with the N-terminal POTRA domains and other YaeT complex members.

Various members of the bacterial-surface antigen superfamily appear to form multimers of four or five independent β-barrels *in vitro*. Purified Tob55 of *N. crassa* predominantly forms a five-ringed multimer of ~15 nm in diameter with a central pore of about 4–5 nm [59]. Purified complexes of the TpsB protein HMW1B in *Haemophilus influenzae* exist mainly as tetramers with a diameter of ~8 nm and a central channel of ~2.5 nm [54]. Although both of these studies demonstrated that the described multimeric structures exhibit pore activity, it is unclear whether it is the individual OMPs within each multimeric structure or the central channel within the ring structures that is responsible for the pore activity (i.e., 4–5 pores per complex vs 1 pore per complex). Structural analysis of Toc75 core complexes by transmission EM indicates that the proteins assemble into structures of four or five OMPs per particle with a diameter of between 12 and 14 nm, and at least in this complex the individual OMPs appear to form independent pores [61]. Interestingly, pore activity measured for reconstituted proteoliposomes of Toc75 and a cyanobacterial homologue of YaeT was very low unless the N-terminal regions of the proteins were absent [58].

Taken together, these data suggest that the β-barrel of YaeT-like proteins are conserved in their ability to form channels in the OM, either independently or when present in a complex, and the N-terminal domains of these proteins may function as a globular plug that gates the activity of the resulting pores [58]. The β-barrel domain of YaeT has also been shown to be sufficient for pore formation [60]. Whether this pore activity is important for function *in vivo* remains to be determined, but, interestingly, the pore activity of YaeT β-barrels appears to fluctuate and may reflect a dynamic and flexible structure that could conceivably accommodate OMP substrates in some form [60].

## X. Conclusions and Future Study

The field of OM biogenesis is undergoing a period of rapid advances with the discovery of novel assembly factors identified using a combination of genetic and biochemical approaches. The discovery of each new protein offers the opportunity to add new pieces of information to what has been the intractable puzzle of envelope biogenesis. In particular, there is renewed promise that some aspects of OMP transport across the bacterial membrane and its subsequent assembly will be better understood given the identification of the YaeT complex and its role in OMP assembly; however, this awaits further elucidation of the functions of each of these components. YaeT-like proteins are a conserved feature responsible for assembling β-barrel proteins in membranes found in organelles as diverse as Gram-negative OMs and eukaryotic mitochondria, but the assembly process in disparate organisms appears to be uniquely fine-tuned as the identity of accessory complex components for a given YaeT-like protein are not necessarily conserved.

Several models of OMP folding and insertion into the OM have been proposed ([48], and reviewed in [60]). For example, the ''folding and lateral release'' model proposes a mechanism analogous to that proposed for transmembrane domains of IM proteins released from the SecYEG translocon [2], whereby YaeT-like porins would undergo major structural rearrangements to reveal a transient lateral opening for release of β-hairpin strands formed within the porin, but it has been argued that this model is thermodynamically and structurally unlikely [48]. A similar hypothesis evoked in reference to Tob55 proposes that the porin may provide an ''Anfinsen-like'' cage environment for OMPs to fold prior to release [59]. The ''scaffold model'' theorizes that Tob55 has an intramembrane chaperone-like function that assists in the otherwise unfavorable β-strand insertion into the OM [44]. This proposal is based on the ''molten disk/globule'' intermediate, a model used to describe distinct kinetic phases measured during spontaneous refolding of OmpA into lipid bilayers *in vitro* and the observation that folding and insertion of OmpA β-hairpin strands into the membrane appear to be coupled processes [62].

Yet another mystery to be resolved is whether OMP folding and insertion into the OM are the last directed steps in the assembly pathway. What prompts the rapid multimerization of trimeric porins? Is this a lipid-catalyzed reaction that occurs subsequent to the release of monomeric OMPs from the assembly machinery, or is some additional protein factor required to initiate trimerization (Figure 6.1)?

We expect that future experiments will broaden our understanding of the functional contributions of the β-barrel and POTRA domains of

YaeT with regard to the folding and assembly of OMP substrates, and should shed some insight into the mechanisms involved in these processes. Furthermore, it will be interesting to elucidate if YaeT shares direct interactions with other factors shown to assist in OMP assembly such as lipids, chaperones, and folding factors [9, 62].

## References


1. Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67:593–656.
2. Dalbey, R.E., and Chen, M. (2004). Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* 1694:37–53.
3. Tokuda, H., and Matsuyama, S. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta* 1693:5–13.
4. Bosch, D., Leunissen, J., Verbakel, J., de Jong, M., van Erp, H., and Tommassen, J. (1986). Periplasmic accumulation of truncated forms of outer-membrane PhoE protein of *Escherichia coli* K-12. *J Mol Biol* 189:449–455.
5. Freudl, R., Schwarz, H., Stierhofg, Y.-D., Gamon, K., Hindennach, I., and Henning, U. (1986). An outer membrane protein (OmpA) of *Escherichia coli* K-12 undergoes a conformational change during export. *J Biol Chem* 261:11355–11361.
6. Werner, J., Augustus, A.M., and Misra, R. (2003). Assembly of TolC, a structurally unique and multifunctional outer membrane protein of *Escherichia coli* K-12. *J Bacteriol* 185:6540–6547.
7. Metcalfe, M., and Holland, I.B. (1980). Synthesis of a major outer membrane porin by *E. coli* spheroplasts. *FEMS Microbiol Lett* 7:111–114.
8. Sen, K., and Nikaido, H. (1990). *In vitro* trimerization of OmpF porin secreted by spheroplasts of *Escherichia coli*. *Proc Natl Acad Sci USA* 87:743–747.
9. Duguay, A.R., and Silhavy, T.J. (2004). Quality control in the bacterial periplasm. *Biochim Biophys Acta* 1694:121–134.
10. Raivio, T.L. (2005). Envelope stress responses and Gram-negative bacterial pathogenesis. *Mol Microbiol* 56:1119–1128.
11. Bayer, M.E. (1968). Adsorption of bacteriophages to adhesions between wall and membrane of *Escherichia coli*. *J Virol* 2:346–356.
12. Muhlradt, P.F., Menzel, J., Golecki, J.R., and Speth, V. (1973). Outer membrane of salmonella. Sites of export of newly synthesised lipopolysaccharide on the bacterial surface. *Eur J Biochem* 35:471–481.
13. Smit, J., and Nikaido, H. (1978). Outer membrane of gram-negative bacteria. XVIII. Electron microscopic studies on porin insertion sites and growth of cell surface of *Salmonella typhimurium*. *J Bacteriol* 135:687–702.
14. Boyd, A., and Holland, I.B. (1980). Intermediate location in the assembly of the matrix protein or porin into the outer membrane of *Escherichia coli*. *J Bacteriol* 143:1538–1541.
15. Kellenberger, E. (1990). The 'Bayer bridges' confronted with results from improved electron microscopy methods. *Mol Microbiol* 4:697–705.
16. Tefsen, B., Geurtsen, J., Beckers, F., Tommassen, J., and de Cock, H. (2005). Lipopolysaccharide transport to the bacterial outer membrane in spheroplasts. *J Biol Chem* 280:4504–4509.

17. Oka, T., and Mihara, K. (2005). A railroad switch in mitochondrial protein import. *Mol Cell* 18:145–146.
18. Koraimann, G. (2003). Lytic transglycosylases in macromolecular transport systems of Gram-negative bacteria. *Cell Mol Life Sci* 60:2371–2391.
19. Sen, K., and Nikaido, H. (1991). Lipopolysaccharide structure required for *in vitro* trimerization of *Escherichia coli* OmpF porin. *J Bacteriol* 173:926–928.
20. Surrey, T., and Jähnig, F. (1992). Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc Natl Acad Sci USA* 89:7457–7461.
21. Bulieris, P.V., Behrens, S., Holst, O., and Kleinschmidt, J.H. (2003). Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *J Biol Chem* 278:9092–9099.
22. De Las Peñas, A., Connolly, L., and Gross, C.A. (1997). $\sigma^E$ is an essential sigma factor in *Escherichia coli*. *J Bacteriol* 179:6862–6864.
23. Mecsas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J., and Gross, C.A. (1993). The activity of $\sigma^E$, an *Escherichia coli* heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev* 7:2618–2628.
24. Sampson, B.A., Misra, R., and Benson, S.A. (1989). Identification and characterization of a new gene of *Escherichia coli* K-12 involved in outer membrane permeability. *Genetics* 122:491–501.
25. Dartigalongue, C., Missiakas, D., and Raina, S. (2001). Characterization of the *Escherichia coli* $\sigma^E$ regulon. *J Biol Chem* 276:20866–20875.
26. Lazar, S.W., and Kolter, R. (1996). SurA assists the folding of *Escherichia coli* outer membrane proteins. *J Bacteriol* 178:1770–1775.
27. Rouviere, P.E., and Gross, C.A. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev* 10:3170–3182.
28. Braun, M., and Silhavy, T.J. (2002). Imp/OstA is required for cell envelope biogenesis in *Escherichia coli*. *Mol Microbiol* 45:1289–1302.
29. Bos, M.P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004). Identification of a lipopolysaccharide transporter in the bacterial outer membrane. *Proc Natl Acad Sci USA* 101:9417–9422.
30. Kabir, M.S., Yamashita, D., Koyama, S., Oshima, T., Kurokawa, K., Maeda, M., Tsunedomi, R., Murata, M., Wada, C., Mori, H., and Yamada, M. (2005). Cell lysis directed by $\sigma^E$ in early stationary phase and effect of induction of the *rpoE* gene on global gene expression in *Escherichia coli*. *Microbiology* 151:2721–2735.
31. Rhodius, V.A., Suh, W.C., Nonaka, G., West, J., and Gross, C.A. (2006). Conserved and variable functions of the $\sigma^E$ stress response in related genomes. *PLoS Biol* 4:e2.
32. Tam, C., and Missiakas, D. (2005). Changes in lipopolysaccharide structure induce the $\sigma^E$-dependent response of *Escherichia coli*. *Mol Microbiol* 55:1403–1412.
33. Takase, I., Ishino, F., Wachi, M., Kamata, H., Doi, M., Asoh, S., Matsuzawa, H., Ohta, T., and Matsuhashi, M. (1987). Genes encoding two lipoproteins in the *leuS-dacA* region of the *Escherichia coli* chromosome. *J Bacteriol* 169:5692–5699.
34. Wu, T., McCandlish, A.C., Gronenberg, L.S., Chng, S.-S., Silhavy, T.J., and Kahne, D. (2006). Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci USA* 103:11754–11759.
35. Eggert, U.S., Ruiz, N., Falcone, B.V., Branstrom, A.A., Goldman, R.C., Silhavy, T.J., and Kahne, D. (2001). Genetic basis for activity differences between vancomycin and glycolipid derivatives of vancomycin. *Science* 294:361–364.
36. Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T.J. (2005). Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* 121:307–317.

37. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J., and Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* 121:235–245.
38. Onufryk, C., Crouch, M.-L., Fang, F.C., and Gross, C.A. (2005). A characterization of six lipoproteins in the $\sigma^E$ regulon. *J Bacteriol* 187:4552–4561.
39. Reumann, S., Davila-Aponte, J., and Keegstra, K. (1999). The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. *Proc Natl Acad Sci USA* 96:784–789.
40. Thomas, K.L., Leduc, I., Olsen, B., Thomas, C.E., Cameron, D.W., and Elkins, C. (2001). Cloning, overexpression, purification, and immunobiology of an 85-kilodalton outer membrane protein from *Haemophilus ducreyi*. *Infect Immun* 69:4438–4446.
41. Genevrois, S., Steeghs, L., Roholl, P., Letesson, J.-J., and van der Ley, P. (2003). The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *EMBO J* 22:1780–1791.
42. Voulhoux, R., Bos, M.P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299:262–265.
43. Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balazsi, G., Ravasz, E., Daugherty, M.D., Somera, A.L., Kyrpides, N.C., Anderson, I., Gelfand, M.S., Bhattacharya, A., Kapatral, V., *et al.* (2003). Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol* 185:5673–5684.
44. Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J Cell Biol* 164:19–24.
45. Cavalier-Smith, T. (2006). Rooting the tree of life by transition analyses. *Biol Direct* 1:19.
46. Bölter, B., Soll, J., Schulz, A., Hinnah, S., and Wagner, R. (1998). Origin of a chloroplast protein importer. *Proc Natl Acad Sci USA* 95:15831–15836.
47. Steeghs, L., den Hartog, R., den Boer, A., Zomer, B., Roholl, P., and van der Ley, P. (1998). Meningitis bacterium is viable without endotoxin. *Nature* 392:449–450.
48. Gentle, I.E., Burri, L., and Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. *Mol Microbiol* 58:1216–1225.
49. Doerrler, W.T., and Raetz, C.R. (2005). Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J Biol Chem* 280:27679–27687.
50. Werner, J., and Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. *Mol Microbiol* 57:1450–1459.
51. Ruiz, N., Wu, T., Kahne, D., and Silhavy, T.J. (2006). Probing the barrier function of the outer membrane with chemical conditionality. *ACS Chem Biol* 1:385–395.
52. Malinverni, J.C., Werner, J., Kim, S., Sklar, J.G., Kahne, D., Misra, R., and Silhavy, T.J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* 61:151–164.
53. Jacob-Dubuisson, F., Fernandez, R., and Coutte, L. (2004). Protein secretion through autotransporter and two-partner pathways. *Biochim Biophys Acta* 1694:235–257.
54. Surana, N.K., Grass, S., Hardy, G.G., Li, H., Thanassi, D.G., and St. Geme, J.W., 3rd (2004). Evidence for conservation of architecture and physical properties of Omp85-like proteins throughout evolution. *Proc Natl Acad Sci USA* 101:14497–14502.
55. Sánchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A. (2003). POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem Sci* 28:523–526.

56. Errington, J., Daniel, R.A., and Scheffers, D.J. (2003). Cytokinesis in bacteria. *Microbiol Mol Biol Rev* 67:52–65.
57. Robson, S.A., and King, G.F. (2006). Domain architecture and structure of the bacterial cell division protein DivIB. *Proc Natl Acad Sci USA* 103:6700–6707.
58. Ertel, F., Mirus, O., Bredemeier, R., Moslavac, S., Becker, T., and Schleiff, E. (2005). The evolutionarily related β-barrel polypeptide transporters from *Pisum sativum* and *Nostoc* PCC7120 contain two distinct functional domains. *J Biol Chem* 280:28281–28290.
59. Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of β-barrel membrane proteins. *Nature* 426:862–866.
60. Stegmeier, J.F., and Andersen, C. (2006). Characterization of pores formed by YaeT (Omp85) from *Escherichia coli*. *J Biochem (Tokyo)* 140:275–283.
61. Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003). Characterization of the translocon of the outer envelope of chloroplasts. *J Cell Biol* 160:541–551.
62. Tamm, L.K., Hong, H., and Liang, B. (2004). Folding and assembly of β-barrel membrane proteins. *Biochim Biophys Acta* 1666:250–263.

This page intentionally left blank

# 7

# *The Function of the ABC Transporter LolCDE in Protein Transport to the Outer Membrane of* E. coli

SHIN-ICHIRO NARITA • HAJIME TOKUDA

*Institute of Molecular and Cellular Biosciences*
*University of Tokyo*
*1-1-1 Yayoi*
*Bunkyo-ku*
*Tokyo 113-0032, Japan*

## I. Abstract

Bacterial lipoproteins represent a subset of membrane proteins possessing a lipid-attached Cys at the N-terminus and are involved in various activities in the cell envelope. Lipoprotein precursors are translocated across the inner (cytoplasmic) membrane by a Sec translocon and processed to the mature forms on the outer surface of the inner membrane. Mature lipoproteins are anchored to membranes through their N-terminal lipid moieties. The Lol system, which is composed of five Lol proteins, is conserved in many Gram-negative bacteria and catalyzes the sorting of hydrophobic lipoproteins through the hydrophilic periplasm to the outer membrane depending on the lipoprotein-sorting signal. The ABC transporter LolCDE complex releases outer membrane lipoproteins from the inner membrane as a lipoprotein–LolA complex. LolA is the periplasmic carrier protein and transports lipoproteins across the periplasm to the outer membrane. LolB in the outer membrane accepts lipoproteins from LolA and then incorporates them into the outer membrane. Only the reaction



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25007-3

catalyzed by the LolCDE complex depends on the lipoprotein-sorting signal and requires ATP. The Asp residue at the second position of lipoproteins functions as a LolCDE avoidance signal, lipoproteins thereby being retained in the inner membrane, whereas others direct lipoproteins to the outer membrane. The crystal structures of LolA and LolB are similar to each other, while their amino acid sequences are dissimilar. Here, we discuss the mechanisms underlying the sorting signal-dependent membrane detachment of lipoproteins by the LolCDE complex and subsequent transport of lipoproteins to the outer membrane.

## II. Introduction

### A. Structure and Function of Outer Membrane Proteins in Gram-Negative Bacteria

The cell envelope of Gram-negative bacteria consists of the inner membrane, the periplasm-containing peptidoglycans, and the outer membrane. The inner membrane is a symmetric phospholipid bilayer, while the outer membrane is an asymmetric bilayer possessing lipopolysaccharides (LPS) in the outer leaflet and phospholipids in the inner leaflet (Figure 7.1). This asymmetric bilayer contains a few major protein species and a number of lipoproteins, which have a lipid-attached Cys at the N-terminus. The major proteins span the outer membrane whereas most lipoproteins are anchored

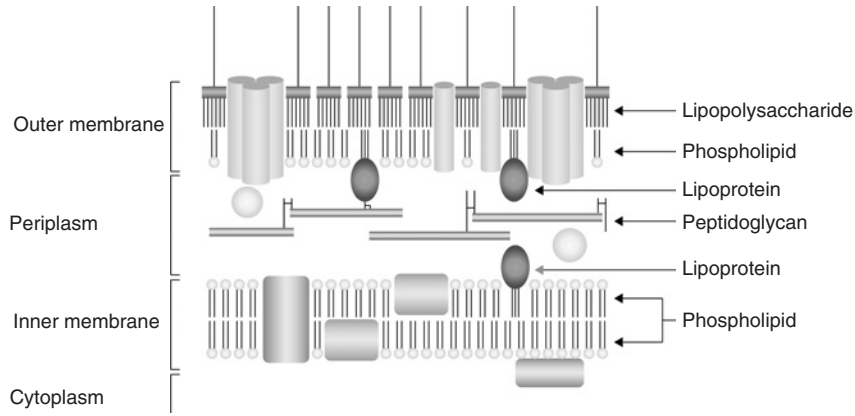



Fig. 7.1. Structure of the *E. coli* envelope. The outer membrane is an asymmetrical bilayer with LPS in the outer leaflet and phospholipids in the inner leaflet. In *E. coli*, lipoproteins are anchored to the periplasmic surface of either the inner or outer membrane.

to the outer membrane through the attached lipids. These proteins do not have hydrophobic stretches that form a transmembrane α-helix and function as stop transfer or signal anchor sequences [1]. Instead, the major outer membrane proteins span the membrane through amphipathic β-strands possessing alternating hydrophobic residues, which do not cause the retention of proteins in the inner membrane. Modification of lipoproteins with lipids occurs on the outer surface of the inner membrane [2] and therefore does not affect translocation of the protein moiety [3]. Thus, both the β structure and lipid modification are characteristic of outer membrane-associated proteins.

Out of 4289 open reading frames (ORFs) in the chromosome of the *Escherichia coli* K12 strain, 898 are predicted to encode inner membrane proteins [4]. According to the *E. coli* cell envelope protein data collection (ECCE, http://www.cf.ac.uk/biosi/staff/ehrmann/tools/ecce/ecce.htm), 343 ORFs of the K12 strain are predicted to encode precursor proteins with signal peptides. The database of bacterial lipoproteins (DOLOP, http://www.mrc-lmb.cam.ac.uk/genomes/dolop/) predicts 86 putative lipoprotein genes in the K12 strain genome [5], 50 ORFs of which are not listed in ECCE. Furthermore, more than 100 ORFs have been predicted to encode lipoproteins by others [6, 7]. Although it is not yet easy to predict β-barrel proteins from their primary sequences, 118 ORFs are predicted to encode integral β-barrel outer membrane proteins [8]. Based on these predictions, the outer membrane seems to have about 100 minor β-barrel proteins, a few major β-barrel proteins, and about 100 lipoproteins, while about 200 proteins are localized in the periplasm.

Many lipoprotein genes are also predicted in other bacteria. For example, 114 are predicted for Gram-positive *Bacillus subtilis* [9] and 105 for Lyme disease spirochaete *Borrelia burgdorferi* [10]. Lipoproteins in these bacteria also seem to be responsible for various membrane-associated activities. However, most lipoproteins including even those of *E. coli* have no known function and are only predicted to be lipoproteins.

Biochemical examination of the proteins encoded by more than 100 putative lipoprotein genes revealed that *E. coli* possesses at least 90 lipoprotein species [11, 12]. Disruption of each of the 90 lipoprotein genes revealed that only two lipoproteins, LolB [13] and YfiO [14], were essential, as reported. However, the disruption of many other genes made cells temperature-sensitive, cold-sensitive, or hypersensitive to drugs. It has been proposed that three lipoproteins, YfiO, YfgL, and NlpB, form a complex with YaeT, which is an *E. coli* homologue of Omp85 [15, 16] and are involved in outer membrane insertion of β-barrel proteins [17]. Indeed, disruption of *yfgL* caused the disappearance of some proteins from the outer membrane. The protein moieties of most *E. coli* lipoproteins are

predicted to be soluble and are presumably exposed to the periplasm, suggesting that they play important roles in various activities in the periplasm.

In *E. coli*, lipoproteins are anchored to the periplasmic side of either the inner or outer membrane depending on the lipoprotein-sorting signal [18]. Some Gram-negative bacteria are known to possess lipoproteins on the outer surface of the outer membrane [19, 20]. In Gram-positive bacteria, lipoproteins are anchored to the outer leaflet of the cytoplasmic membrane.

## B. Biogenesis of Lipoproteins

Lipoproteins are each synthesized in the cytoplasm as a precursor with an N-terminal signal peptide and are then translocated across the inner membrane by the Sec translocon [2]. Lipoproteins that depend on the Tat pathway for translocation across the inner membrane have not been reported, although such lipoproteins may exist [21]. Targeting of *E. coli* lipoproteins to the Sec translocon has not been extensively studied. However, CyoA, BRP, and Lpp have been shown to require the SRP pathway [22, 23], while Lpp also utilizes the SecB pathway for targeting [23].

The maturation of lipoprotein precursors sequentially occurs on the periplasmic surface of the inner membrane (Figure 7.2A). Phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt) recognizes the consensus sequence -Leu-(Ala/Ser)-(Gly/Ala)-Cys-, called the lipobox, around the cleavage site of the signal peptide and transfers the diacylglyceryl moiety from phospatidylglycerol to the sulfhydryl group of Cys of the lipobox to yield a modified prolipoprotein. Signal peptidase II (LspA, also called prolipoprotein signal peptidase) then recognizes the diacylglyceryl cysteine and cleaves the signal peptide, rendering the S-lipidated Cys a new N-terminus. The apolipoprotein thus formed is further modified with an acyl chain at the amino group of the Cys by phospholipid:apolipoprotein transacylase (Lnt) [24]. Mature lipoproteins therefore have an *N*-acyl-*S*-diacylglyceryl-Cys at the N-terminus (Figure 7.2B). Lipoproteins of Gram-positive bacteria are thought to lack the amino-linked acyl chain because the gene for Lnt has not been found in them [2, 9, 25].

Since globomycin, an inhibitor of signal peptidase II, does not inhibit the Sec-dependent translocation of OmpA, diacylglyceryl prolipoproteins seem to be laterally extruded from the Sec translocon into the lipid bilayer. In this case, the lipoprotein signal peptide may be regarded as a signal anchor sequence of integral inner membrane proteins. It has been reported that YidC is required for the targeting and translocation of some lipoproteins [22, 23, 26–28].

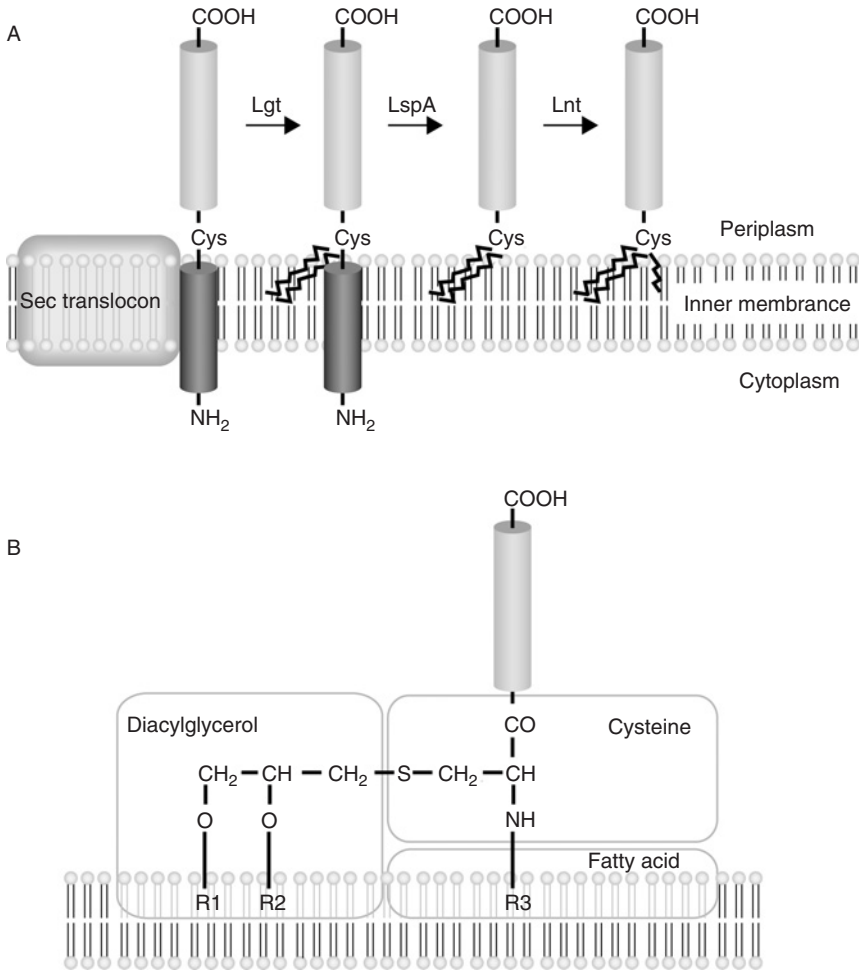



FIG. 7.2. Biogenesis of lipoproteins. (A) Processing of a lipoprotein precursor takes place on the periplasmic side of the inner membrane after it has been translocated across the inner membrane through the action of the Sec translocon. Reactions are sequentially catalyzed by three enzymes; phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt), signal peptidase II (LspA), and phospholipid:apolipoprotein transacylase (Lnt). (B) The N-terminal Cys of mature lipoproteins is modified with thioether-linked diacylglycerol and amino-linked acyl chains.

### C. Lipoprotein-Sorting Signals

Lipoproteins are localized to either the inner or the outer membrane depending on the lipoprotein-sorting signal. Yamaguchi *et al.* [29] first proposed a rule for the membrane specificity of *E. coli* lipoproteins. They found that replacement of Ser at position 2 by Asp caused inner membrane localization of an outer membrane-specific lipoprotein. Furthermore, replacement of Asp at position 2 by another residue caused outer membrane localization of an inner membrane-specific lipoprotein. From these observations, it was proposed that Asp at position 2 functions as an inner membrane retention signal for lipoproteins, while other residues cause outer membrane localization. It was later found that the residue at position 3 affects the inner membrane retention of lipoproteins due to Asp at position 2 [30]. Seydel *et al.* [31] constructed a maltose-binding protein (MalE) derivative having a lipid-attached Cys at the N-terminus (lipoMalE). When the lipoMalE derivative had Asp at position 2, a chromosomal *malE* deletion mutant could grow in the presence of maltose as a sole source of carbon because the derivative was localized in the inner membrane and functioned as a maltose-binding protein. In contrast, the *malE* mutant did not grow when lipoMalE derivatives were localized in the outer membrane. It was then found that lipoMalE derivatives possessing Phe, Trp, Tyr, Gly, or Pro at position 2 also supported mutant growth, indicating that Asp at position 2 is not the sole inner membrane retention signal. The residue at position 3 of lipoMalE was Asn in these experiments. However, it is noteworthy that *E. coli* native lipoproteins do not have Phe, Trp, Tyr, Gly, or Pro at position 2. It has also been reported that the formation of a higher order of conformation also affects the inner membrane retention of lipoproteins due to Asp at position 2 [32].

When the residue at position 3 was Ser, only Asp at position 2 caused the retention of lipoproteins in the inner membrane [33], indicating that Asp at position 2 is the intrinsic inner membrane retention signal. Subsequent comprehensive studies on the effect of residues at position 3 revealed that the potency of Asp at position 2 as the inner membrane retention signal is highest with Asp, Glu, or Gln at position 3 [33]. Importantly, all *E. coli* lipoproteins with Asp at position 2 have one of these three amino acids at position 3 [12], indicating that strong inner membrane signals are utilized by native inner membrane lipoproteins. Asn at position 2 with Asp at position 3 was found in the inner membrane lipoprotein AcrE [34, 35] and also found to function as the inner membrane retention signal [31, 33].

## III. Sorting of Lipoproteins by the Lol System

### A. Localization of Lipoproteins

Lipoproteins are extremely hydrophobic due to their N-terminal lipid moiety. Lipoproteins to be localized in compartments other than the inner membrane must cross the hydrophilic periplasm to reach their final destinations after they have been processed to their mature forms on the inner membrane. *E. coli* lipoproteins are localized on the periplasmic leaflet of either the inner or the outer membrane. On the other hand, pullulanase, PulA, of *Klebsiella pullulanase* is localized on the outer leaflet of the outer membrane. The type II secretion system, which is composed of 15 proteins, specifically mediates the transport of PulA from the periplasm to the outer leaflet of the outer membrane [2]. Some bacteria, such as *Borrelia* species, also have lipoproteins on the outer leaflet of the outer membrane [20, 36]. In this case, however, the mechanism is not known.

The *l*ocalization *o*f *l*ipoproteins (Lol) system, which is the main subject of this chapter, was identified in *E. coli* and found to be conserved in various Gram-negative bacteria. The Lol system composed of five Lol proteins catalyzes the transport of lipoproteins from the outer leaflet of the inner membrane to the inner leaflet of the outer membrane depending on the lipoprotein-sorting signal (Figure 7.3).

### B. LolA, a Periplasmic Chaperon for Lipoproteins

Among the five Lol proteins, the first one identified was LolA, which is required for the release of the major outer membrane lipoprotein Lpp from spheroplasts. When Lpp was expressed in spheroplasts, it was processed to the mature form but not released into the spheroplast supernatant, while nonlipidated outer membrane proteins were released from the spheroplasts. However, Lpp was released into the supernatant when it was expressed in the presence of concentrated periplasmic materials [37]. LolA, which is responsible for the Lpp-releasing activity, was then purified as a 20-kDa protein from the periplasmic fractions and shown to form a soluble complex with Lpp in a molar ratio of 1:1. LolA also caused the release of other outer membrane lipoproteins such as Pal, NlpB, RlpA, RlpB, Slp, LolB [38], and chimeric Lpp-β-lactamase lipoprotein, indicating that LolA is generally required for the release of outer membrane lipoproteins. In contrast, LolA did not cause the release of inner membrane-specific lipoproteins with Asp at position 2. These results revealed that the LolA-dependent release determines whether lipoproteins are localized to the outer membrane or remain in the inner membrane. LolA depletion

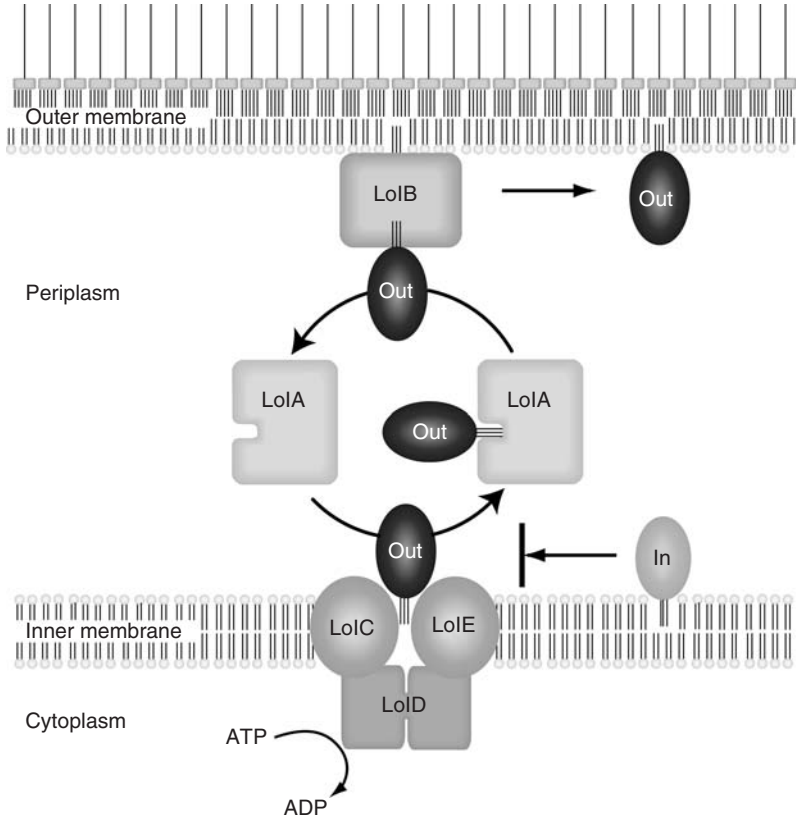



FIG. 7.3. Sorting of lipoproteins to the outer membrane by the Lol system. "In" and "Out" represent the inner and outer membrane lipoproteins, respectively. The ABC transporter LolCDE recognizes outer membrane lipoproteins and releases them from the inner membrane, causing the formation of a complex between one molecule each of lipoprotein and LolA, a periplasmic carrier protein. The LolA–lipoprotein complex traverses the periplasm to the outer membrane, where lipoproteins are transferred to the receptor protein LolB and then incorporated into the outer membrane. LolB is itself an outer membrane lipoprotein. Inner membrane lipoproteins with Asp at position 2 avoid recognition by LolCDE and thus remain anchored to the inner membrane.

caused the accumulation of outer membrane lipoproteins in the inner membrane and severe inhibition of growth [39].

### C. LolB, an Outer Membrane Receptor for Lipoproteins

When the lipoprotein–LolA complex reaches the outer membrane, lipoproteins are dissociated from LolA and incorporated into the outer membrane [37, 40]. The incorporation of lipoproteins into the outer membrane was

inhibited when the outer membrane was pretreated with trypsin, indicating that a proteinaceous factor in the outer membrane is involved in this reaction. This factor was identified and named LolB, which is itself a lipoprotein [40]. However, examination revealed that the membrane anchor was dispensable for the LolB function. A LolB derivative, mLolB, lacking the N-terminal acyl chain-attached Cys accepted lipoproteins from LolA and incorporated them into the outer membrane. Moreover, mLolB also incorporated lipoproteins into the inner membrane and even liposomes (Tsukahara, Narita, and Tokuda, unpublished data). These results indicate that mLolB targets associated lipoproteins to the lipid bilayer and causes their incorporation into it. LolB is essential for *E. coli*, and its depletion causes the accumulation of lipoproteins in the periplasm [13, 40]. However, expression of mLolB enabled the *E. coli* Δ*lolB* strain to grow although the outer membrane-specific-lipoproteins were transiently mislocalized to the inner membrane.

### D. Structures of LolA and LolB

Both LolA and LolB have molecular masses of about 20 kDa and exhibit common functions, that is, binding and transfer of lipoproteins. Despite these similarities, their amino acid sequences are dissimilar. However, the X-ray crystal structures of LolA and mLolB determined at 1.65- and 1.9-Å resolution, respectively, revealed the striking similarity in their three-dimensional structures [41]. Both have a hydrophobic cavity composed of an unclosed β-barrel and an α-helical lid (Figure 7.4). Lipoproteins are highly hydrophobic but become water soluble when they form a complex with LolA or mLolB, suggesting that the lipid moiety of lipoproteins is shielded in this hydrophobic cavity. The LolA lid is closed by the hydrogen bonds between Arg at position 43 in the barrel and residues in the α1- and α2 helices. Therefore, it is speculated that the lid opens when a lipoprotein is accommodated in the hydrophobic cavity of LolA. On the other hand, the hydrophobic cavity of mLolB is always open.

Mutational analyses of LolA revealed residues that are critical for the transfer of lipoproteins to LolB [42–45]. Replacement of Arg at position 43 by Leu did not affect the ability to bind a lipoprotein, but abolished that to transfer a lipoprotein to LolB. Further mutation of this residue revealed that the amino acid residue at position 43 affects the efficiency of the lipoprotein transfer to LolB. Hydrophobic residues at this position increase the hydrophobic interaction between lipoprotein and LolA, thereby inhibiting the transfer of lipoproteins to LolB [44]. It was also found that the hydrophobic interaction between LolA and lipoproteins is considerably weaker than that between LolB and lipoproteins (Figure 7.5). This is critically important for one-way transfer of lipoproteins from LolA to LolB in an energy-independent manner.

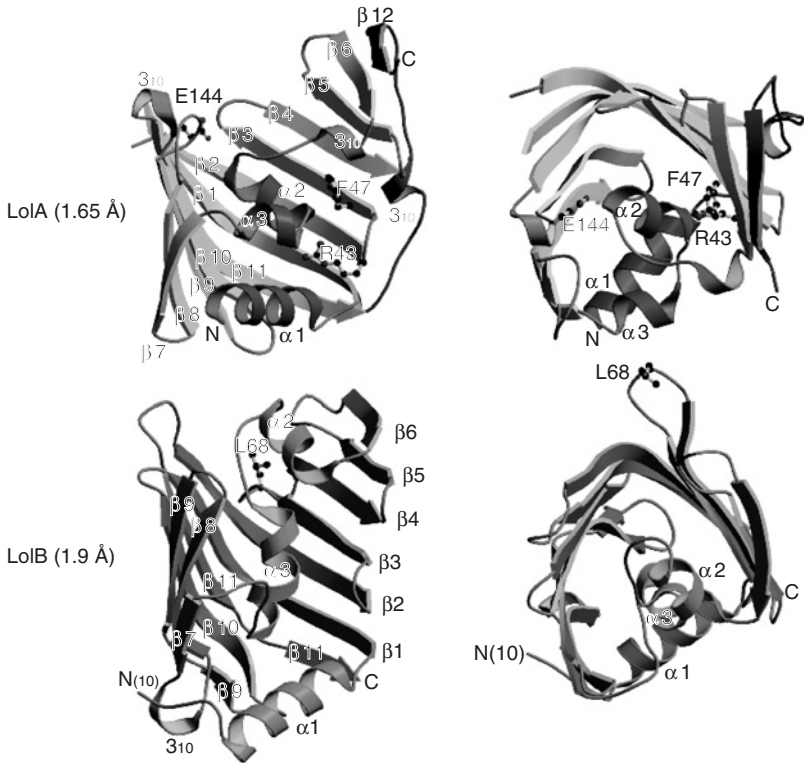



FIG. 7.4. Crystal structures of LolA and LolB. The LolA and LolB molecules are each shown as a ribbon model. The structural information on LolA (1UA8) and LolB (1IWM) was obtained from the RCSB protein data bank (http://pdb.protein.osaka-u.ac.jp/pdb/).

The three-dimensional structure of LolB revealed a novel Leu residue at position 68. This residue is conserved among LolB homologues of the $\gamma$ subdivision of proteobacteria and is located in the loop that protrudes into the solvent region despite its hydrophobic nature [41] (Figure 7.4), suggesting its importance for targeting to the lipid bilayer. Indeed, replacement of Leu at position 68 by Glu or Arg abolished the membrane-targeting function of mLolB without affecting its ability to bind lipoproteins (Tsurumizu *et al*., unpublished data).

## E. LolCDE, an ABC Transporter Mediating the Membrane Detachment of Lipoproteins

### *1. Identification of the Lipoprotein-Releasing Apparatus*

The LolA-dependent release of lipoproteins from right-side out membrane vesicles required ATP [46]. GTP and UTP substituted for ATP while ADP, AMP, or AMP-PNP, a nonhydrolyzable ATP analogue, did not cause

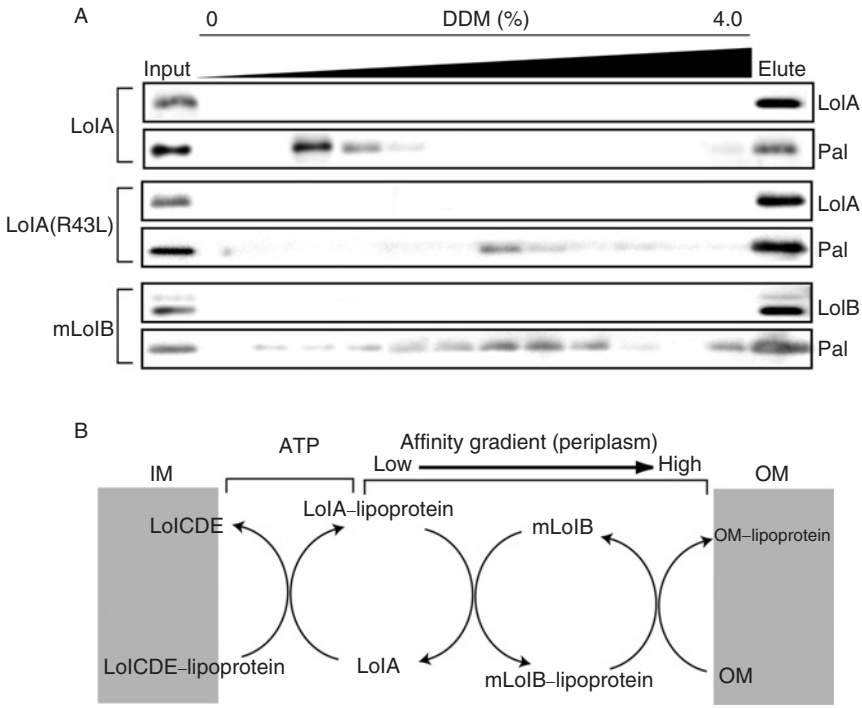



FIG. 7.5. Intensity of the hydrophobic interaction between Lol factors and Pal. (A) Pal was released from spheroplasts as a complex with LolA or the LolA(R43L) derivative and then adsorbed to a TALON column through a His-tag attached to the C-terminus of the LolA protein. The mLolB–Pal complex was prepared by incubating the LolA–Pal complex with His-tagged mLolB and also adsorbed to a TALON. The column was eluted with a linear gradient of *n*-dodecyl-β-D-maltopyranoside (DDM), and fractions were analyzed by SDS-PAGE and immunoblotting with anti-LolA, -Pal, and -LolB antibodies. The left lane shows 50% amounts of samples applied to the resin. A significant amount of Pal became dissociated from the wild-type LolA with a very low concentration of DDM, whereas a higher concentration of DDM was required for the dissociation of Pal from the LolA(R43L) derivative. The concentration of DDM required to dissociate Pal from mLolB was similar to that required for the dissociation from LolA (R43L), suggesting that the intensity of the hydrophobic interaction is similar for the LolA (R43L)-Pal and mLolB–Pal complexes. This is most likely the reason why LolA(R43L) cannot transfer lipoproteins to LolB. (B) Energy-independent transfer of lipoproteins from LolA to LolB to the outer membrane takes place in the direction in which the affinity for lipoproteins increases. Therefore, the LolA–lipoprotein interaction should be weak, otherwise lipoprotein transfer to LolB would be inhibited, causing accumulation of lipoproteins in the periplasm.

the release of lipoproteins. Sodium vanadate completely inhibited the ATP-dependent release of lipoproteins. Based on these observations, solubilized inner membrane proteins and outer membrane-specific lipoprotein Pal were reconstituted into proteoliposomes and then the release of Pal, which was dependent on both LolA and ATP, was examined. The inner membrane

proteins responsible for the lipoprotein release were then purified by successive chromatographies. N-terminal sequencing revealed that one of the proteins purified was the product of *ycfV*, and possessed Walker A and B motifs and an ABC signature motif (Figure 7.6A), the last of which is a characteristic sequence of ABC ATPases [47]. The *ycfV* gene (o228b) is flanked by two genes, *ycfU* and *ycfW*, encoding putative transmembrane proteins, suggesting that the three genes form a single transcriptional unit [3, 48]. The products of these three genes formed a complex and exhibited lipoprotein-releasing activity when reconstituted into proteoliposomes. This complex, designated as LolCDE, was composed of one copy each of

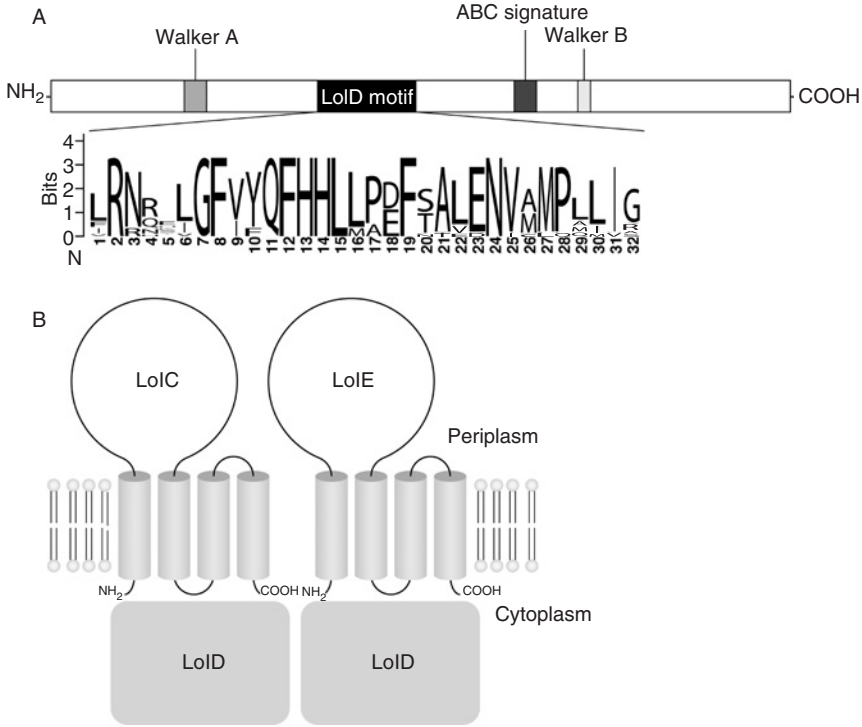



FIG. 7.6. Structure of the LolCDE complex. (A) Schematic representation of conserved regions of LolD. The graphical representation of the LolD motif was generated using WebLogo software [78], by aligning the sequences of orthologues from the α, β, γ and δ subdivisions of Gram-negative proteobacteria. (B) Topology of the LolCDE complex. Both LolC and LolE span the membrane four times, and have a large domain exposed to the periplasm between the first and second transmembrane helices. The membrane topology of LolC and LolE is based on prediction and biochemical analyses (Yasuda *et al*., unpublished data).

transmembrane proteins LolC (YcfU) and LolE (YcfW), and two copies of ATPase LolD (YcfV; Figure 7.6).

ABC transporters, both importers and exporters, mediate the transmembrane movement of substrates. On the other hand, the LolCDE complex mediates the release of lipoproteins after they have been translocated across the inner membrane by the Sec translocon. Defective LolD mutations had no effect on the generation of mature lipoproteins. These results indicate that LolCDE is involved in neither the translocation of lipoprotein precursors across the inner membrane nor the formation of mature lipoproteins [3]. *E. coli* has been predicted to possess 57 ABC transporters [48]. So far as reported, only two ABC transporters, LolCDE [49] and MsbA [50], are essential for the growth of *E. coli*. MsbA has been proposed to catalyze the transmembrane movement of LPS [51, 52]. In addition to LolA and LolB, LolC, LolD, and LolE are essential for *E. coli* growth. Depletion of LolCDE did not affect the translocation of lipoprotein precursors across the inner membrane but inhibited the release of lipoproteins [49]. These results indicate that the LolCDE complex is the sole apparatus mediating the release of *E. coli* lipoproteins.

ABC transporters generally have more than 10 transmembrane segments (TMS) [47, 53]. In contrast, both LolC and LolE span the membrane four times with a large domain exposed to the periplasm (Figure 7.6B; Yasuda *et al.*, unpublished data). Therefore, the total number of TMS in the LolCDE complex is eight, which is less than that in most other ABC transporters. This may be related to the unique function of the LolCDE complex, which catalyzes the release of substrates from the outer leaflet of the lipid bilayer.

### 2. *Effect of Sorting Signals on the Recognition of Lipoproteins by LolCDE*

Proteoliposomes reconstituted with the LolCDE complex did not release Pal(S2D), which is a derivative of Pal and has Asp in place of Ser at position 2, indicating that lipoprotein-sorting signals function at the release step [3]. Furthermore, these results also indicate that no factor other than LolA and the LolCDE complex is required for recognition of the sorting signals. The ATPase activity of the LolCDE complex was stimulated by Pal but not by Pal(S2D) [54]. Moreover, the release of Lpp by the LolCDE complex from proteoliposomes was competitively inhibited by an excess amount of Pal, whereas an excess amount of Pal(S2D) had no effect [54]. Taken together, these results revealed an intriguing property of the inner membrane signal, that is, Asp at position 2 functions as the LolCDE avoidance signal, causing retention of lipoproteins in the inner membrane (Figure 7.7).

ATP hydrolysis by ABC transporters is generally stimulated by their substrates, although the extent of stimulation varies with the transporter.

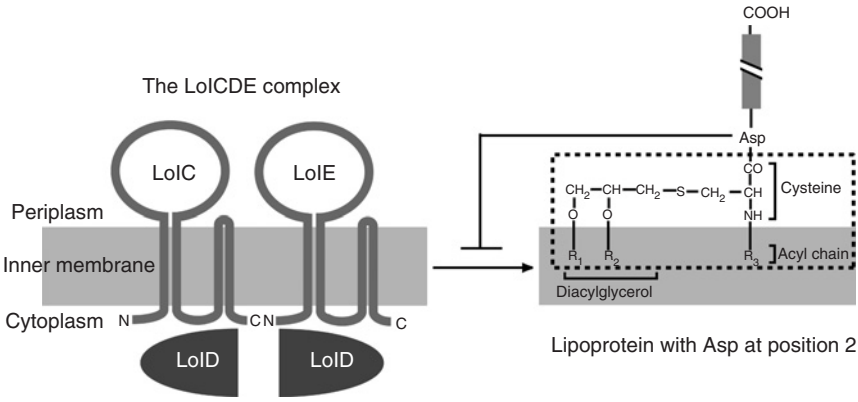



FIG. 7.7. The LolCDE avoidance function of Asp at position 2 of lipoproteins. LolCDE recognizes the N-terminal Cys together with the attached acyl chains (broken square) but not the second residue. When Asp is present at position 2, it inhibits the recognition of lipoproteins by LolCDE, thereby causing the retention of lipoproteins in the inner membrane. The phospholipid composition is critically important for the LolCDE avoidance mechanism.

Stimulation of the ATPase activity of LolCDE by Pal was observed in the absence of LolA. Moreover, a leaky LolCDE mutant causing the mislocalization of inner membrane lipoproteins to the outer membrane has been isolated [55]. This mutant carried an Ala to Pro mutation at position 40 of LolC. Both LolA and LolB interacted with inner membrane lipoproteins released through the action of this LolC mutant, causing the localization of inner membrane lipoproteins in the outer membrane. These results, taken together, indicate that lipoprotein-sorting signals only function at the LolCDE complex.

An apolipoprotein lacking the amino-linked acyl chain was not released from proteoliposomes [56], indicating that the N-acylation of lipoproteins is essential for the recognition by the LolCDE complex. Depletion of Lnt *in vivo* also resulted in the accumulation of apolipoproteins in the inner membrane [57]. However, it should be pointed out that the mode of lipid modification at the N-terminal Cys is the same whether lipoproteins are specific to the inner or the outer membrane [56], indicating that the lipoprotein-sorting signal does not affect lipid modification but functions as a determinant of membrane localization.

To further understand the functions of the residue at position 2, Pal(S2C) possessing Cys in place of Ser at position 2 was subjected to Cys-specific chemical modification, and the subsequent release reaction was examined in reconstituted proteoliposomes. The release of Pal(S2C) was not affected by modification with large nonprotein molecules such as iodoacetyl biotin [58],

indicating that LolCDE releases lipoproteins without recognizing the structure of the residue at position 2 (Figure 7.7). Then the question arises; how Asp at position 2 functions as the LolCDE avoidance signal and causes the retention of lipoproteins in the inner membrane. The answer is most likely the participation of membrane phospholipids in the avoidance mechanism. Cys at position 2 of Pal(S2C) in proteoliposomes was not modified with membrane-impermeable SH-specific reagent, indicating that the residue at position 2 of lipoproteins is not exposed to the solvent [58]. *E. coli* contains phosphatidylethanolamine (PE), phosphatidylglycerol, and cardiolipin (CL) as major phospholipids. When proteoliposomes contained a mixture of these three phospholipids, the LolCDE complex released lipoproteins in a sorting signal-dependent manner. On the other hand, chemical modification of PE with an amine-specific reagent, sulfo-*N*-hydroxysuccinimide acetate, resulted in the release of not only Pal but also Pal(S2D). When proteoliposomes were reconstituted with CL alone, Pal (S2D) was released in LolCDE- and LolA-dependent manners [58]. When Asp at position 2 was modified with a carboxylate-specific reagent, the avoidance function was abolished even in proteoliposomes reconstituted with *E. coli* phospholipids [58]. Taken together, these results suggest that the positive charge of PE and the negative charge of Asp are required for the avoidance function, presumably for electrostatic interaction. Although Glu at position 2 does not function as an inner membrane signal [30, 33], oxidation of Cys at position 2 with performic acid to cysteic acid caused generation of the LolCDE avoidance signal. The distance between Cα and the negative charge of Asp is 2.72–3.69 Å, which is similar to that in the case of cysteic acid (2.87–3.95 Å), while the distance for Glu is 4.27–4.93 Å [58]. These results indicate that the Lol avoidance signal should have a negative charge that is within a certain distance from Cα of the second residue. From these observations, it has been proposed that the steric and electrostatic interactions between Asp at position 2 of lipoproteins and PE are critical for the LolCDE avoidance mechanism. It is speculated that the N-terminal conformation of inner membrane lipoproteins is distinct from that of outer membrane lipoproteins due to these interactions, inhibition of the recognition by LolCDE thereby being caused [58]. The A40P mutation of LolC mentioned above may alter the conformation of the LolCDE complex so that it can accept the unique conformation of the N-terminal region of inner membrane lipoproteins.

PE is known as a nonbilayer lipid [59] and has been proposed to be important for the functions of various membrane proteins [60–63]. CL also forms nonbilayer structures in the presence of a high concentration of magnesium ions [64]. It was found that the phospholipid composition affects not only the Lol avoidance function of Asp at position 2 but also the activity

of the LolCDE complex. Both ATP hydrolysis and the release of lipoproteins by LolCDE reconstituted into CL proteoliposomes significantly increased with an increase in the concentration of $Mg^{2+}$. However, the Lol avoidance function of Asp was not seen in the CL proteoliposomes and inner membrane-specific lipoproteins were efficiently released in the presence of a high concentration of $Mg^{2+}$ (Miyamoto and Tokuda, unpublished data). These results suggest that the nonbilayer structure of phospholipids enhances the activity of LolCDE, while the positive charge of phospholipids plays an important role in lipoprotein sorting. Therefore, PE has dual functions: one is the specific interaction with Asp at position 2 of inner membrane lipoproteins, and the other is the stimulation of the LolCDE activity. In contrast, CL in the presence of a high concentration of $Mg^{2+}$ has only the latter function. An *E. coli* mutant that cannot synthesize PE has been isolated. This mutant requires ~50-mM $Mg^{2+}$ for growth and contains CL and PG, each at about 50% [65]. However, the localization of lipoproteins in this mutant is normal even though PE is completely absent (Miyamoto *et al.*, unpublished data). It was then found that phosphatidylglycerol causes the retention of inner membrane lipoproteins through preferential inhibition of the release of lipoproteins with Asp at position 2.

### 3. *Role of ATP Energy in the Transport of Outer Membrane Lipoproteins*

The only step that requires exogenous energy input is the lipoprotein release reaction catalyzed by LolCDE in the presence of LolA. This step most likely involves the opening of the LolA lid, which was revealed by X-ray crystallographic analysis [41]. For this, the energy obtained through ATP hydrolysis by LolD on the cytoplasmic side of the inner membrane should be transferred to LolC and/or LolE, and then utilized for both the membrane detachment of lipoproteins and opening of the LolA lid. Communication between LolD and LolC/LolE is therefore critically important but scarcely understood.

The sequence between the Walker A and ABC signature motifs is highly conserved among LolD homologues of proteobacteria but not other ABC transporters of *E. coli* [3]. This sequence, the LolD motif, comprises 32 residues (Figure 7.6A) and has been speculated to be a contact site with LolC/E based on the crystal structures of other ABC transporters [66, 67]. Targeted random mutagenesis of these 32 residues resulted in the isolation of dominant-negative mutants, whose overexpression arrested the growth of *E. coli* cells despite the chromosomal *lolD*$^+$ background. Some mutations lowered the ATPase activity of the LolCDE complex with little effect on the ATPase activity of the LolD subunit, suggesting that these mutations perturbed the communication between the membrane-spanning subunits and the

ATPase subunit. Indeed, the dominant-negative phenotype of these LolD mutants could be suppressed by secondary mutations in LolC or LolE [68]. Interestingly, mutations of the LolC suppressors were mainly located in the periplasmic loop whereas ones of LolE suppressors were mainly located in the cytoplasmic loop, suggesting that the mode of interaction with LolD differs between LolC and LolE.

### *4. Closer Look at the Molecular Events in the Initial Sorting Reactions*

When the LolCDE complex was purified in the absence of ATP, various outer membrane lipoproteins, but not inner membrane ones, were copurified. In contrast, when ATP was present during the purification procedure, no lipoproteins were associated with the purified LolCDE complex (Ito *et al*., unpublished data). These results indicate that the liganded LolCDE complex represents the intermediate of the lipoprotein release reaction formed in the inner membrane. LolCDE is the first example of an ABC transporter that was isolated with tightly bound substrates. ATP caused dissociation of lipoproteins from the liganded LolCDE in a detergent-dependent manner, indicating that the interaction between LolCDE and lipoproteins is hydrophobic. ATP hydrolysis was not required for this dissociation. The $K_m$ for ATP was higher with liganded LolCDE than with unliganded LolCDE, whereas they exhibited similar $V_{max}$ values. Based on these observations, detailed molecular events occurring in the inner membrane are speculated to be as follows: (1) The sorting cycle is initiated by binding of outer membrane-specific lipoproteins to LolC/E in an ATP-independent manner, resulting in an increase in the affinity of LolD for ATP. Liganded LolCDE can be purified at this stage. (2) ATP binding to LolD causes a decrease in the affinity of LolC/E for lipoproteins through a conformational change. Unliganded LolCDE can be purified at this stage. (3) ATP hydrolysis induces membrane detachment of lipoproteins and formation of the LolA–lipoprotein complex. This step most likely involves opening of the LolA lid for the accommodation of lipoproteins in its hydrophobic cavity, as discussed earlier. It is difficult to maintain the amount of lipoproteins in proteoliposomes sufficiently high enough for a continuous release reaction. Therefore, the ATPase activity of the LolCDE complex, which is tightly coupled to the release reaction, remains to be determined.

### *5. Evolutionary Conservation of the Lol System*

Five Lol factors, LolA through LolE, are highly conserved in the $\gamma$ subdivision of proteobacteria. Moreover, bacteria in other subdivisions also possess Lol factors, although some bacteria lack the gene for LolE.

The lipoprotein-releasing ABC transporters in these bacteria most likely each function as a dimer of the LolC-LolD heterodimer. It seems certain that the Lol system is conserved among Gram-negative bacteria and functions to sort a number of lipoproteins to outer membranes. Sequences similar to the LolD motif are also found in the ABC proteins of Gram-positive bacteria such as Actinomycetes. The *salX* gene of *Streptococcus salivarius* encodes the ABC subunit of the export machinery for lantibiotic. SalX exhibits moderate sequence similarity with LolD throughout its sequence, including the region that corresponds to the LolD motif. It should be pointed out that some LolD homologues of Gram-negative bacteria are incorrectly designated as antimicrobial peptide transporters. The MJ0796 protein of *Methanococcus jannaschii*, an archaea, is also a LolD homologue and was crystallized [69]. The structure of this protein provides valuable information, although it is not clear whether it exhibits any functional relation to LolD.

Gram-negative bacteria, such as *Brucella suis* and *Desulfovibrio vulgaris*, apparently lack the gene for LolB, while they possess those for LolC, LolD, and LolA. The outer membrane lipoproteins of these bacteria may be released from the inner membrane through a mechanism homologous to that of *E. coli*, but may be incorporated into outer membranes through a different mechanism. It may be possible that a certain outer membrane protein or lipoprotein has a dissimilar sequence to LolB but retains the LolB function. Alternatively, LolA of these bacteria may also have the LolB function.

Despite conservation of the Lol system, the sorting signals of lipoproteins may not always be conserved in Gram-negative bacteria, in which more than 2000 putative lipoprotein genes have been found [5]. Inner membrane lipoproteins that have residues other than Asp are found in many bacteria. The chromosome and plasmids of spirochaete *B. burgdorferi* encode more than 130 potential lipoproteins, many of which are assumed to be located on the outer surface of the outer membrane [10, 70]. It was found that neither the inner membrane localization of OppAIV nor the outer surface localization of OspA is determined by the residue at position 2 [20]. Moreover, it was proposed through the use of a fluorescent reporter protein, that *Borrelia* lipoproteins are localized on the outer surface by default, but are retained on the inner membrane through an unknown mechanism. *Neisseria meningitides*, which belongs to the β subdivision of proteobacteria, has a complete set of Lol factors. Two of the three homologues of *Neisserial* DsbA are lipoproteins and are involved in the formation of disulfide linkages in the periplasm. The one has Ser and the other Asp at position 2, while both lipoproteins are localized in the inner membrane [71]. Therefore, the residue at position 2 does not seem

to determine the inner membrane localization of these DsbA homologues. AcrA is an inner membrane lipoprotein with Asp at position 2 and constitutes a multidrug-efflux pump with AcrB and TolC in *E. coli*. However, the AcrA homologue of *Pseudomonas aeruginosa*, MexA, has Gly at position 2. Detailed *in vivo* analyses revealed that Lys at position 3 and Ser at position 4, but not Gly at position 2, are critical for the inner membrane-retention of MexA (Narita and Tokuda, unpublished data), although these two residues are not conserved among paralogues of MexA. Lipoprotein-sorting signals in *P. aeruginosa* were also examined in proteoliposomes reconstituted with *P. aeruginosa* LolCDE in the presence of *P. aeruginosa* LolA. Lys and Ser at positions 3 and 4, respectively, were found to retain lipoproteins in proteoliposomes. Interestingly, Asp at position 2 also functioned as an inner membrane retention signal (Tanaka, Narita, and Tokuda, unpublished data). Taken together, these observations indicate that Asp at position 2 is a general LolCDE avoidance signal, which is independent of the origin of Lol factors, while the Lys-Ser residues at positions 3 and 4 are species-specific inner membrane retention signals. It remains to be clarified whether or not the Lys-Ser signals also function as a LolCDE avoidance signal in *P. aeruginosa*. Taken together, the mechanism for the sorting of lipoproteins seems to be more diverse and complex than previously proposed [58].

## F. Perspectives

Lipid modification of proteins is an effective strategy by which even hydrophilic proteins can be localized in membranes. Replacement of the signal peptides of some outer membrane lipoproteins, such as CsgG and Wza, with nonlipoprotein type signal peptides inhibited their correct assembly and functions [72, 73]. On the other hand, membrane localization by the lipid anchor is dispensable for the functions of some lipoproteins [74, 75], whereas the lipid anchor is important for the efficiency of correct localization of other lipoproteins [76, 77]. As far as the Lol system is concerned, the N-terminal Cys with attached acyl chains is essential for the recognition by LolCDE (Figure 7.7). The mode of recognition of lipoproteins by the LolCDE complex has been extensively studied in the last few years. It has been revealed in *E. coli* that residues other than Asp at position 2 have no effect on the recognition of lipoproteins by LolCDE. Asp at position 2 is considered to interact with phospholipids and to disturb recognition of the N-terminal Cys by the LolCDE complex. However, the interaction between the Asp residue and phospholipids has not been directly proved. In addition to the effect on the Lol avoidance function of Asp, the phospholipid composition was found to significantly affect the

ATPase and lipoprotein-releasing activities of LolCDE. Moreover, lipoprotein-sorting signals were found to be diverse in Gram-negative bacteria even though the Lol system is widely conserved. To understand the general mechanism by which bacterial lipoproteins are sorted, detailed analyses of the LolCDE complex of various bacteria are required.

## Acknowledgments

We wish to thank Rika Ishihara for the help in the preparation of this chapter. The work described here was supported by grants to H.T. from the Ministry of Education, Science, Sports, and Culture of Japan.

## References

1. von Heijne, G. (1996). Principles of membrane protein assembly and structure. *Prog Biophys Mol Biol* 66:113–139.
2. Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* 57:50–108.
3. Yakushi, T., Masuda, K., Narita, S., Matsuyama, S., and Tokuda, H. (2000). A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat Cell Biol* 2:212–218.
4. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580.
5. Madan Babu, M., Leena Priya, M., Tamli Selvan, A., Madera, M., Gough, J., Aravind, L., and Sankaran, K. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol* 188:2761–2773.
6. Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 12:1652–1662.
7. Brokx, S.J., Ellison, M., Locke, T., Bottorff, D., Frost, L., and Weiner, J.H. (2004). Genome-wide analysis of lipoprotein expression in *Escherichia coli* MG1655. *J Bacteriol* 186:3254–3258.
8. Zhai, Y., and Saier, M.H., Jr. (2002). The β-barrel finder (BBF) program, allowing identification of outer membrane b-barrel proteins encoded within prokaryotic genomes. *Protein Sci* 11:2196–2207.
9. Tjalsma, H., and van Dijl, J.M. (2005). Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics* 17:4472–4482.
10. Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gwinn, M., Dougherty, B., *et al.* (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* 390:580–586.
11. Miyadai, H., Tanaka-Masuda, K., Matsuyama, S., and Tokuda, H. (2004). Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. *J Biol Chem* 279:39807–39813.

12. Tokuda, H., Matsuyama, S., and Tanaka-Masuda, K. (2006). Structure, function and transport of lipoproteins in *Escherichia coli*, in M. Ehrmann (ed.), *Periplasm*, pp. 69–79. Washington, DC: ASM Press.
13. Tanaka, K., Matsuyama, S., and Tokuda, H. (2001). Deletion of *lolB,* encoding an outer membrane lipoprotein, is lethal for *Escherichia coli* and causes the accumulation of lipoprotein localization intermediates in the periplasm. *J Bacteriol* 183:6538–6542.
14. Onufryk, C., Crouch, M.-L., Fang, F.C., and Gross, C.A. (2005). Characterization of six lipoproteins in the $\sigma^E$ regulon. *J Bacteriol* 187:4552–4561.
15. Genevrois, S., Steeghs, L., Roholl, P., Letesson, J.J., and van der Ley, P. (2003). The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *EMBO J* 22:1780–1789.
16. Voulhoux, R., Bos, M.P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299:262–265.
17. Wu, T., Malinverni, J., Ruiz, N., Kim, S., and Silhavy, T.J. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* 121:235–245.
18. Tokuda, H., and Matsuyama, S. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta* 1693:5–13.
19. Haake, D.A. (2000). Spirochaetal lipoproteins and pathogenesis. *Microbiology* 146: 1491–1504.
20. Schulze, R.J., and Zückert, W.R. (2006). *Borrelia burgdorferi* lipoproteins are secreted to the outer surface by default. *Mol Microbiol* 59:1473–1484.
21. Sargent, F., Berks, B.C., and Palmer, T. (2006). Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* 254:198–207.
22. Bloois, E.V., Haan, G.J., Gier, J.W., Oudega, B., and Luirink, J. (2006). Distinct requirements for translocation of the N-tail and C-tail of the *Escherichia coli* inner membrane protein CyoA. *J Biol Chem* 281:10002–10009.
23. Fröderberg, L., Houben, E.N., Baars, L., Luirink, J., and de Gier, J.W. (2004). Targeting and translocation of two lipoproteins in *Escherichia coli* via the SRP/Sec/YidC pathway. *J Biol Chem* 279:31026–31032.
24. Sankaran, K., and Wu, H.C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J Biol Chem* 269:19701–19706.
25. Sutcliffe, I.C., and Harrington, D.J. (2002). Pattern searches for the identification of putative lipoprotein genes in gram-positive bacterial genomes. *Microbiology* 148:2065–2077.
26. Samuelson, J.C., Jiang, F., Yi, L., Chen, M., de Gier, J.-W., Kuhn, A., and Dalbey, R.E. (2001). Function of YidC for the insertion of M13 procoat protein in *Escherichia coli*. *J Biol Chem* 276:34847–34852.
27. du Plessis, D.J., Nouwen, N., and Driessen, A.J. (2006). Subunit A of cytochrome o oxidase requires both YidC and SecYEG for membrane insertion. *J Biol Chem* 281:12248–12252.
28. Celebi, N., Yi, L., Facey, S.J., Kuhn, A., and Dalbey, R.E. (2006). Membrane biogenesis of subunit II of cytochrome bo oxidase: contrasting requirements for insertion of N-terminal and C-terminal domains. *J Mol Biol* 357:1428–1436.
29. Yamaguchi, K., Yu, F., and Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* 53:423–432.
30. Gennity, J.M., and Inouye, M. (1991). The protein sequence responsible for lipoprotein membrane localization in *Escherichia coli* exhibits remarkable specificity. *J Biol Chem* 266:16458–16464.

31. Seydel, A., Gounon, P., and Pugsley, A.P. (1999). Testing the '2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol Microbiol* 34:810–821.
32. Robichon, C., Bonhivers, M., and Pugsley, A.P. (2003). An intramolecular disulphide bond reduces the efficacy of a lipoprotein plasma membrane sorting signal. *Mol Microbiol* 49:1145–1154.
33. Terada, M., Kuroda, T., Matsuyama, S., and Tokuda, H. (2001). Lipoprotein sorting signals evaluated as the LolA-dependent release of lipoproteins from the inner membrane of *Escherichia coli*. *J Biol Chem* 276:47690–47694.
34. Klein, J.R., Henrich, B., and Plapp, R. (1991). Molecular analysis and nucleotide sequence of *envCD* operon of *Escherichia coli*. *Mol Gen Genet* 230:230–240.
35. Seiffer, D., Klein, J.R., and Plapp, R. (1993). EnvC, a new lipoprotein of cytoplasmic membrane of *Escherichia coli*. *FEMS Microbiol Lett* 107:175–178.
36. Brandt, M.E., Riley, B.S., Radolf, J.D., and Norgard, M.V. (1990). Immunogenic integral membrane proteins of *Borrelia burgdorferi* are lipoproteins. *Infect Immun* 58:983–991.
37. Matsuyama, S., Tajima, T., and Tokuda, H. (1995). A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. *EMBO J* 14:3365–3372.
38. Yokota, N., Kuroda, T., Matsuyama, S., and Tokuda, H. (1999). Characterization of the LolA-LolB system as the general lipoprotein localization mechanism of *Escherichia coli*. *J Biol Chem* 274:30995–30999.
39. Tajima, T., Yokota, N., Matsuyama, S., and Tokuda, H. (1998). Genetic analyses of the *in vivo* function of LolA, a periplasmic chaperone involved in the outer membrane localization of *Escherichia coli lipoproteins*. *FEBS Lett* 439:51–54.
40. Matsuyama, S., Yokota, N., and Tokuda, H. (1997). A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. *EMBO J* 16:6947–6955.
41. Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S., Tokuda, H., and Miki, K. (2003). Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. *EMBO J* 22:3199–3209.
42. Miyamoto, A., Matsuyama, S., and Tokuda, H. (2001). Mutant of LolA, a lipoprotein-specific molecular chaperone of *Escherichia coli*, defective in the transfer of lipoproteins to LolB. *Biochem Biophys Res Commun* 287:1125–1128.
43. Miyamoto, A., Matsuyama, S., and Tokuda, H. (2002). Dominant-negative mutant of a lipoprotein-specific molecular chaperone, LolA, tightly associates with LolCDE. *FEBS Lett* 528:193–196.
44. Taniguchi, N., Matsuyama, S., and Tokuda, H. (2005). Mechanism underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed β-barrel structures. *J Biol Chem* 280:34481–34488.
45. Watanabe, S., Matsuyama, S., and Tokuda, H. (2006). Roles of the hydrophobic cavity and lid of LolA in the lipoprotein transfer reaction in *Escherichia coli*. *J Biol Chem* 281:3335–3342.
46. Yakushi, T., Yokota, N., Matsuyama, S., and Tokuda, H. (1998). LolA-dependent release of a lipid-modified protein from the inner membrane of *Escherichia coli* requires nucleotide triphospate. *J Biol Chem* 273:32576–32581.
47. Holland, I.B., and Blight, M.A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* 293:381–399.
48. Linton, K.J., and Higgins, C.F. (1998). The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Mol Microbiol* 28:5–13.

49. Narita, S., Tanaka, K., Matsuyama, S., and Tokuda, H. (2002). Disruption of *lolCDE*, encoding an ATP-binding-cassette transporter, is lethal for *Escherichia coli* and prevents the release of lipoproteins from the inner membrane. *J Bacteriol* 184:1417–1422.
50. Doerrler, W.T., Reedy, M.C., and Raetz, C.R.H. (2001). An *Escherichia coli* mutant defective in lipid export. *J Biol Chem* 276:11461–11464.
51. Doerrler, W.T., Gibbons, H.S., and Raetz, C.R.H. (2004). MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J Biol Chem* 279:45102–45109.
52. Tefsen, B., Bos, M.P., Beckers, F., Tommassen, J., and de Cock, H. (2005). MsbA is not required for phospholipid transport in *Neisseria meningitidis*. *J Biol Chem* 280:35961–35966.
53. Biemans-Oldehinkel, E., Doeven, M.K., and Poolman, B. (2006). ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett* 580:1023–1035.
54. Masuda, K., Matsuyama, S., and Tokuda, H. (2002). Elucidation of the function of lipoprotein-sorting signals that determine membrane localization. *Proc Natl Acad Sci USA* 99:7390–7395.
55. Narita, S., Kanamaru, K., Matsuyama, S., and Tokuda, H. (2003). A mutation in the membrane subunit of an ABC transporter LolCDE complex causing outer membrane localization of lipoproteins against their inner membrane-specific signals. *Mol Microbiol* 49:167–177.
56. Fukuda, A., Matsuyama, S., Hara, T., Nakayama, J., Nagasawa, H., and Tokuda, H. (2002). Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. *J Biol Chem* 277:43512–43518.
57. Robichon, C., Vidal-Ingigliardi, D., and Pugsley, A.P. (2005). Depletion of apolipoprotein *N*-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. *J Biol Chem* 280:974–983.
58. Hara, T., Matsuyama, S., and Tokuda, H. (2003). Mechanism underlying the inner membrane retention of *E. coli* lipoproteins caused by Lol avoidance signals. *J Biol Chem* 278:40408–40414.
59. Gruner, S. (1985). Intrinsic curvature hypothesis for biomembrane composition: a role for nonbilayer lipids. *Proc Natl Acad Sci USA* 82:3665–3669.
60. Curnow, P., Lorch, M., Charalambous, K., and Booth, P.J. (2004). The reconstitution and activity of the small multidrug transporter EmrE is modulated by non-bilayer lipid composition. *J Mol Biol* 343:213–222.
61. Booth, P.J., Curran, A.R., Templer, R.H., Lu, H., and Meijberg, W. (2001). Manipulating the folding of membrane proteins: using the bilayer to our advantage. *Biochem Soc Symp* 68:27–33.
62. van der Does, C., Swaving, J., van Klompenburg, W., and Driessen, A.J. (2000). Non-bilayer lipids stimulate the activity of the reconstituted bacterial protein translocase. *J Biol Chem* 275:2472–2478.
63. Mikhail, B., Sun, J., Kaback, H.R., and Dowhan, W. (1996). A phospholipid acts as a chaperone in assembly of a membrane transport protein. *J Biol Chem* 271:11615–11618.
64. Rietveld, A.G., Killian, J.A., Dowhan, W., and de Kruijff, B. (1993). Polymorphic regulation of membrane phospholipid composition in *Escherichia coli*. *J Biol Chem* 268:12427–12433.
65. DeChavigny, A., Heacock, P.N., and Dowhan, W. (1991). Sequence and inactivation of the *pss* gene of *Escherichia coli*. Phosphatidylethanolamine may not be essential for cell viability. *J Biol Chem* 266:5323–5332.
66. Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* 296:1091–1098.

67. Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J., and Hunt, J.F. (2002). ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimmer. *Mol Cell* 10:139–149.
68. Ito, Y., Matsuzawa, H., Matsuyama, S., Narita, S., and Tokuda, H. (2006). Genetic analysis of the mode of interplay between an ATPase subunit and membrane subunits of the lipoprotein-releasing ATP-binding cassette transporter LolCDE. *J Bacteriol* 188: 2856–2864.
69. Yuan, Y.-R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P.J., and Hunt, J.F. (2001). The crystal structure of the MJ0796 ATP-binding cassette. *J Biol Chem* 276:32313–32321.
70. Casjens, S., Palmer, N., van Vugt, R., Huang, W.M., Stevenson, B., Rosa, P., Lathigra, R., Sutton, G., Peterson, J., Dodson, R.J., Haft, D., Hickey, E., *et al.* (2000). A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol Microbiol* 35:490–516.
71. Tinsley, C.R., Voulhoux, R., Betetti, J.L., Tommassen, J., and Nassif, X. (2004). Three homologues, including two membrane-bound proteins, of the disulfide oxidoreductase DsbA in *Neisseria meningitidis*: effects on bacterial growth and biogenesis of functional type IV pili. *J Biol Chem* 279:27078–27087.
72. Robinson, L.S., Ashman, E.M., Hultgren, S.J., and Chapman, M.R. (2006). Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol Microbiol* 59:870–881.
73. Nesper, J., Hill, C.M., Paiment, A., Harauz, G., Beis, K., Naismith, J.H., and Whitfield, C. (2003). Translocation of group 1 capsular polysaccharide in *Escherichia coli* serotype K30. Structural and functional analysis of the outer membrane lipoprotein Wza. *J Biol Chem* 278:49763–49772.
74. Zgurskaya, H.I., and Nikaido, H. (1999). AcrA is a highly asymmetric protein capable of spanning the periplasm. *J Mol Biol* 285:409–420.
75. Beck, B.J., and Downs, D.M. (1999). A periplasmic location is essential for the role of the ApbE lipoprotein in thiamine synthesis in *Salmonella typhimurium*. *J Bacteriol* 181:7285–7290.
76. Nakajima, A., Sugimoto, Y., Yoneyama, H., and Nakae, T. (2000). Localization of the outer membrane subunit OprM of resistance-nodulation-cell division family multicomponent efflux pump in *Pseudomonas aeruginosa*. *J Biol Chem* 275:30064–30068.
77. Poquet, I., Faucher, D., and Pugsley, A.P. (1993). Stable periplasmic secretion intermediate in the general secretory pathway of *Escherichia coli*. *EMBO J* 12:271–278.
78. Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190.

# Part II

## *Crossing Endoplasmic Reticulum Membranes*

This page intentionally left blank

# 8

# *The Signal Recognition Particle and Its Receptor in ER Protein Targeting*

IAIN L. MAINPRIZE • FELICIA VULCU • DAVID W. ANDREWS

*Department of Biochemistry and Biomedical Sciences*
*McMaster University*
*Hamilton, Ontario, Canada L8N 3Z5*

## I. Abstract

The complex mechanisms that regulate cotranslational targeting of proteins to the endoplasmic reticulum (ER) begin with the translation of a signal sequence in the nascent protein that encodes the required targeting information. In eukaryotes, the recognition and correct sorting of signal sequence-containing polypeptides involves a relatively complicated seven-component ribonucleoprotein complex termed signal recognition particle (SRP) and an ER membrane-bound SRP receptor (SR), a heterodimer composed of two unusual GTPases. The orchestration of these and other components involved in the targeting of proteins to the ER are described in this chapter. The discovery and elucidation of this pathway is presented from a historical perspective that focuses on what is known about SRP and SR. However, to fully explain cotranslational protein targeting requires analysis of the signal hypothesis, mRNA partitioning and alternate mechanisms regulating targeting of the nascent polypeptide chain. Attempts to analyze the structure of the components involved have met with considerable success and the challenge ahead involves making use of this data to extend our understanding of the activity and composition of the ER-directed protein targeting machinery.



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25008-5

## II. Introduction

The process of targeting secretory and membrane proteins to the endoplasmic reticulum (ER) membrane of eukaryotic organisms utilizes a well-established ubiquitous pathway known as the signal recognition particle (SRP) pathway (Figure 8.1). During the synthesis in the cytoplasm of a protein destined for the ER, the ribosome synthesizing it is sorted away from ribosomes synthesizing other proteins when an N-terminal hydrophobic signal sequence emerges from the ribosome exit tunnel and is bound by the SRP. The SRP-bound complex of the ribosome and the associated nascent chain, collectively referred to as the ribosome-nascent chain complex (RNC), can then be targeted to the ER.

The eukaryotic SRP is a ribonucleoprotein complex composed of six polypeptides bound to a 300-nucleotide molecule of RNA, referred to as SRP RNA. The binding of SRP to the signal sequence induces a pause in, or slowing of, translation that maintains the RNC in a translocation-competent state thereby increasing the time available for it to be targeted to the ER membrane. At the membrane, SRP binds the RNC to translocation sites on the ER by interacting with the SRP receptor (SR). SR is composed of two polypeptides, SR$\alpha$, a peripheral membrane component, and SR$\beta$, a Type I (i.e., N-terminus in the ER lumen) transmembrane protein. The concerted interaction between SRP and SR is orchestrated by GTP-binding and hydrolysis cycles; both polypeptides of the SR and the 54-kDa polypeptide of SRP (SRP54) contain GTP-binding sites. Through a still uncharacterized process, the binding of SRP to SR facilitates the docking of the RNC onto the translocation complex (the translocon) such that continued translation of the nascent chain results in transport of the nascent protein chain through the Sec61 complex that forms the translocation pore within the translocon. The nascent chain is either inserted into or translocated across the ER membrane by means of specific interactions with components of the translocon.

Here, we describe what is currently known about the mechanism(s) used to segregate cotranslationally proteins that are targeted to the ER from those that remain in the cytoplasm. This process begins with initiation of translation and ends once the RNC has functionally engaged the translocation pore at the ER membrane. We provide a historical context for the view of these mechanisms and describe areas of controversy. We have emphasized investigations aimed at understanding structure–function relationships for the components of the SRP cycle, as this is currently an area of intense activity. The subsequent steps that transport the nascent protein into or across the ER membrane are covered in Chapter 10. Some proteins that are synthesized entirely in the

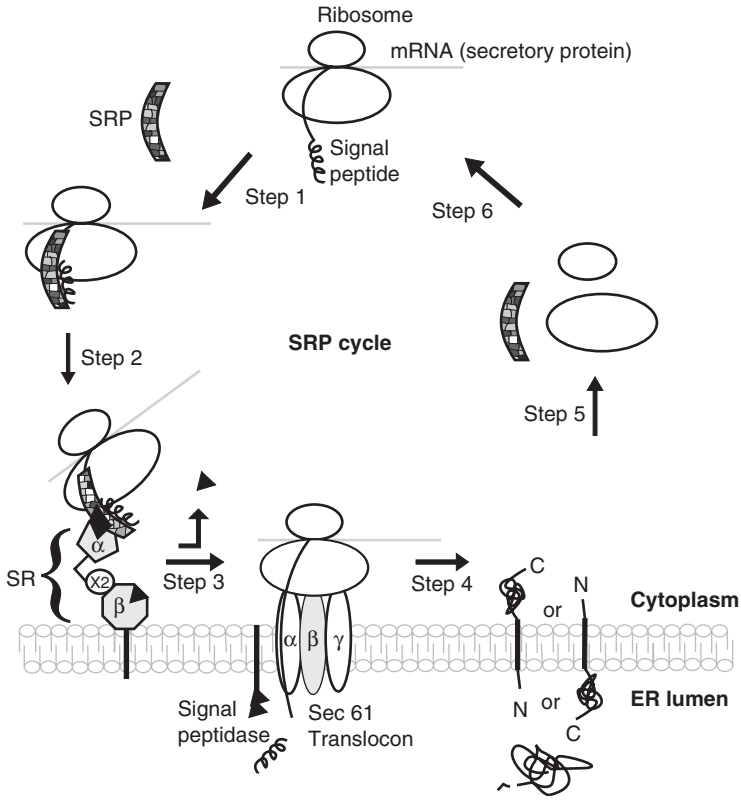



FIG. 8.1. The SRP pathway. Translation of secretory proteins is initiated in the cytoplasm with the binding of the ribosome to the mRNA of a secretory protein. On emergence of the N-terminal hydrophobic signal sequence, the ribonucleoprotein complex SRP binds (Step 1), induces a temporary translational pause, and targets the resulting RNC to the ER membrane where SRP interacts with its receptor, the SR (Step 2). The interaction between SR and SRP is GTPase driven as SRP54, SRα, and SRβ all contain GTPase domains. This process is regulated by an as-of-yet unresolved GTP cycle (guanine nucleotides are indicated as solid triangles). The interaction of SRP and SR results in the subsequent transfer of the RNC to the Sec61 translocon (Step 3), which mediates the full translocation through or integration into the ER membrane (Step 4) to yield a number of protein topologies. The signal peptide is cleaved by signal peptidase during translation and translocation across the membrane. The three topologies shown here are Type I transmembrane proteins (N-terminus in the lumen), Type II transmembrane proteins (N-terminus in the cytosol), and a fully secreted proteins (lumen bound). Once translation terminates, the ribosome can be reused or, like SRP, it can be released from the membrane (Step 5) and be recycled (Step 6).

cytoplasm can be targeted posttranslationally to the ER and a wide variety of other subcellular locations, but they appear to use targeting mechanisms distinct from the cotranslational mechanisms described here ([1], and Chapters 14–18).

## III. Cotranslational Translocation: A Historical Perspective

The first indication that proteins were transported across the membrane of the ER cotranslationally came from the 1966 observations made by Sabatini and Redman when they examined the fate of polypeptides synthesized in an *in vitro* system made from liver microsomes in which the aminoglycoside antibiotic, puromycin, was added to release prematurely nascent polypeptides from the ribosomes [2]. Puromycin displaces the aminoacyl-tRNA from the acceptor site of the ribosome and becomes incorporated at the C-terminus of the growing polypeptide, releasing the nascent chain from the ribosome. Unexpectedly, the released peptidyl-puromycin polypeptides were found associated with the microsomal membranes rather than in the supernatant fraction. Furthermore, these truncated polypeptides could only be dissociated into the soluble fraction by treatment with detergent or mild sonication indicating that they were enclosed within the microsomal lumen [2].

This and other evidence led to the proposal of a mechanism that exists in the cell to target polypeptides to the ER membrane while synthesis is occurring in the ribosome. It was also suggested that unidirectional movement of polypeptides through a cavity composed of the ribosome peptide exit tunnel coupled to a tunnel that spans the entire ER membrane results in vectorial transport into the ER cisternal space [2].

Subsequently, ribosomes were observed to remain attached to the ER membrane even after the nascent polypeptide was released by puromycin addition. However, these ribosomes could be removed from the ER membrane with increasing ionic strength, a phenomenon that would not occur prior to puromycin treatment [3]. This result demonstrated that there is a receptor that anchors the ribosome to the ER membrane and suggested that cytosolic and membrane-bound ribosomes are structurally and functionally interchangeable. Additionally, the finding that ribosomes remained bound to the ER after premature release of the nascent polypeptide led to the postulate that the nature of the messenger RNA (mRNA) dictates the eventual fate of the cognate polypeptide. Certain aspects of this mRNA-partitioning model remain relevant to our understanding of the mechanisms responsible for targeting of proteins to the ER and are discussed in more detail below.

## A. The Signal Hypothesis

The discovery in the 1970s that secretory proteins destined for the ER carry a common feature in their N-termini, termed the signal sequence, largely displaced mRNA partitioning as a viable mechanism for targeting nascent polypeptides to the ER membrane. A typical signal sequence is 20–25 amino acids long and composed of a core hydrophobic region of 7–15 amino acids and two flanking regions, an N-terminal region that has a net positive charge, and a polar C-region that contains the signal peptidase cleavage site [4, 5]. Even though in 1971 the detailed sequence characteristics of signal peptides were unknown, Sabatini and Blobel first proposed the signal hypothesis based on observations suggesting that signal peptide cleavage correlated with targeting to the ER membrane (Figure 8.1). Although proposed for the targeting of proteins to the ER, polypeptide encoded targeting signals have proven to be a general phenomenon and are responsible for targeting proteins to most intracellular locations. The recognition of the ER-specific signal sequence and the subsequent targeting of the polypeptide to the ER are both mediated by SRP and its cognate receptor (SR) that is situated on the membrane, as described earlier. Translocation across the membrane was proposed to occur via a proteinaceous pore [6]. These two aspects of the signal hypothesis, that the signal sequence directs transport of the nascent polypeptide to the membrane and that translocation occurs through a proteinaceous pore, are now not easily separated. However, for many years after the signal hypothesis was first proposed, alternative theories were common in which targeting and translocation depended on a hydrophobic signal sequence but were driven by thermodynamics and therefore, largely independent of other proteins.

The reconstitution of cotranslational translocation across the ER in 1975 using functional rough microsomes from dog pancreas incubated with ribosomal subunits derived from free ribosomes and programmed with immunoglobulin mRNAs provided the first formal experimental support for the details of the signal hypothesis. A protease protection assay was used to show the successful cotranslational translocation of immunoglobulins into the ER lumen (where they were protected from added protease unless detergent was added to solubilize the ER membrane). Translocation occurred concomitant with cleavage of the signal sequence [7] by what was later identified as the membrane-bound signal peptidase (Figure 8.1).

The first demonstration of one of the central tenets of the signal hypothesis, that the signal sequence is both necessary and sufficient to target a protein to the ER, came from the development of *in vitro* translation systems that could be programmed with RNAs made by *in vitro* transcription of recombinant plasmids encoding fusion proteins. Using this kind of

system, the signal sequence of the prokaryotic periplasmic protein β-lactamase was fused to the normally cytoplasmic protein, α-globin. The chimeric protein was translocated across the ER membrane cotranslationally when expressed in a rabbit reticulocyte cell-free system containing canine rough microsomes [8]. This result not only demonstrated that a signal peptide was sufficient to mediate cotranslational translocation for an otherwise cytoplasmic protein but also demonstrated that the essential features of signal peptides are conserved from *Escherichia coli* to mammals.

The successful translocation of chimpanzee globin using a lactamase signal peptide also highlighted another mystery that remains unresolved. Signal peptides show very limited sequence similarity yet appear to be largely interchangeable. Although some evidence indicates that signal peptides are optimized to their specific passenger [9], the effect on targeting efficiency can be small suggesting that optimization of the signal to match the mature portion of the protein may have been a minor contributor to signal sequence diversity. There is also some evidence that signal peptides may have other cellular functions that are manifest only after cleavage from the mature polypeptide [10]. Such functions may have coevolved with adaptations to the machinery that increased tolerance of diversity or may have arisen because tolerance existed within the system.

The combination of recombinant DNA technology and cell-free translation systems permit unprecedented probing of the substrate requirements for targeting to the ER membrane. For example, by fusing 109 amino acids of the α-globin sequence at the N-terminus of preprolactin (a protein with an N-terminal signal sequence) it was possible to demonstrate that the signal sequence does not have to be located at the N-terminus of a nascent protein. Importantly, the internal signal sequence mediated translocation of the N-terminal globin domain demonstrating that both the targeting and translocation machinery can process a relatively large and, presumably, partially folded protein sequence [11]. Moreover, cell-free systems were used to demonstrate that transmembrane sequences can also be targeted to the ER membrane via the SRP pathway [12]. It was subsequently shown that both Type I (N-terminus lumenal) and Type II (N-terminus cytoplasmic) transmembrane proteins can be targeted to membranes cotranslationally by signal-anchor sequences that interact with most of the same protein components bound by signal peptides. These sequences combine the targeting function of the signal peptide with the membrane integration activity of a transmembrane sequence. As they ultimately reside in the bilayer of the ER membrane, these sequences are not cleaved from the nascent membrane protein and therefore do not interact with signal peptidase.

There is one report that showed that, in multispanning membrane proteins, the additional transmembrane domains that follow a signal anchor do not

require SRP to mediate targeting to membranes [13]. However, the SRP independence of subsequent transmembrane sequences may not be true for all proteins (e.g., ones with a long cytoplasmic domain between transmembrane sequences). Furthermore, it is not clear whether only the initial ribosome on a polysome needs to be targeted via the SRP pathway. A report demonstrating that SRP and SR may be unnecessary when mRNAs are translated by ribosomes already bound to functional translocons suggests that this area requires further investigation [14]. Thus, although it is clear that a variety of sequences with different ultimate fates are targeted to the ER membrane by a conserved process, many of the details are known for only simple substrates [15, 16]. As it is very likely that cotranslational targeting is also a regulatory step in protein biogenesis [15, 16], future discoveries in this area will likely come from examining more exotic substrates. Efficient cell-free systems will be useful not only to provide further insight into the mechanisms of protein targeting but will likely continue to be essential for the identification of additional components that regulate the SRP-mediated targeting pathway (Figure 8.1).

In 1981, Walter *et al.* used a cell-free reticulocyte lysate translation system supplemented with $^{125}$I-labeled SRP to show direct binding of SRP to *in vitro* assembled polysomes synthesizing the secretory protein preprolactin. This binding was specific as SRP did not interact with assembled polysomes synthesizing cytoplasmic proteins like the α- and β-chain of rabbit globin [17]. It was later observed that the targeting efficiency dropped abruptly for polypeptide chains with more than 140 amino acids in front of the signal peptide indicating that the affinity of SRP drops when the nascent chain reaches a certain length. This led to the proposal that SRP scans RNCs preferentially during the initial phase of translation [18]. An attractive feature of the model is that it partially accounts for how SRP can efficiently target ER-destined RNCs even though there are many more ribosomes in the cytoplasm than SRP. However, it does not explain efficient targeting of RNCs encoding proteins like SREBP in which the first sequence that could interact with SRP is located 477 amino acids from the N-terminus of the protein [19].

## B. mRNA Partitioning

The signal hypothesis has also been used to account for mRNA partitioning to ER membranes. Targeting of the RNC would be expected to result in the accumulation of mRNAs encoding secreted and integral membrane proteins at the ER (Figure 8.2). However, if the SRP pathway is responsible for all ER-specific protein targeting then how is it that inactivation of the SRP pathway is not a lethal event [20, 21]? Loss of the SRP

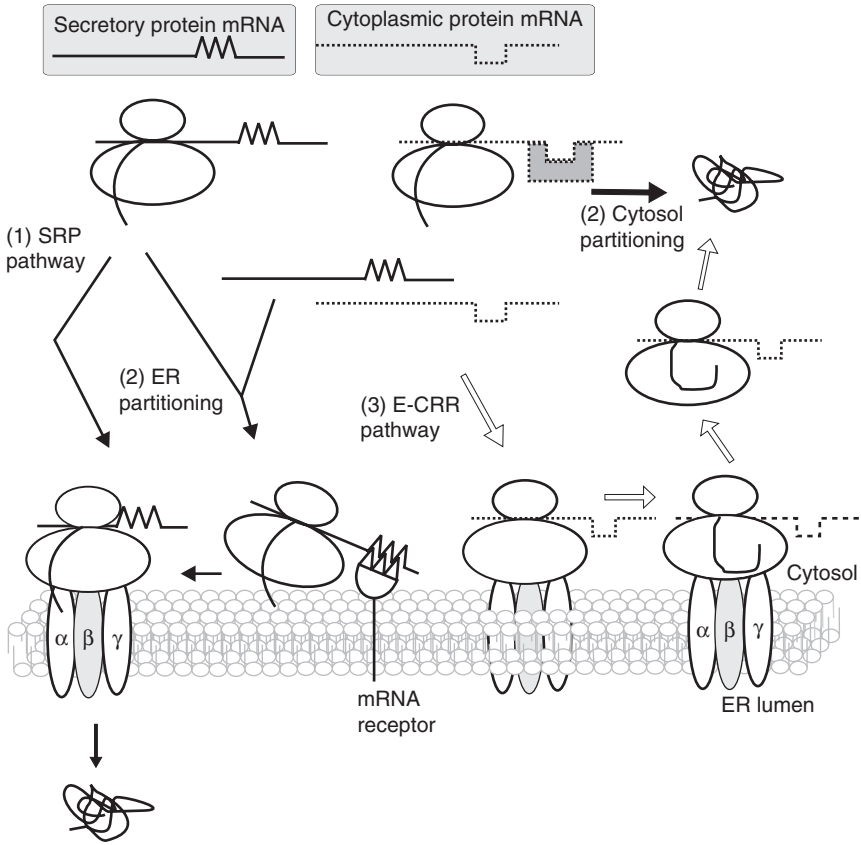



FIG. 8.2. Pathways for mRNA partitioning. (1) The SRP pathway (described in Figure 8.1) partitions secretory proteins by binding of SRP to the N-terminal signal sequence and targeting the resulting RNC to the ER membrane whereby the RNC is docked onto the Sec61 translocon through the actions of the SR. (2) The template-partitioning model hypothesizes the presence of protein-binding regions on the mRNA of secretory or cytosoplasmic proteins that function to sequester these mRNAs to the appropriate compartments. On the ER, the mRNA-binding protein would function as a "secretory" mRNA receptor. Cytosol partitioning of mRNA would be mediated by a protein (shown bound to a cytoplasmic mRNA) that would prevent the mRNA from interacting with membranes. (3) The E-CRR pathway proposes the initiation of translation of cytosolic mRNAs on membrane-bound ribosomes with the subsequent liberation of these ribosomes into the cytosoplasm shortly after translation initiation due to lack of signals to maintain these ribosomes on the membrane. Translation of these mRNAs begins on the membrane but is completed in the cytoplasm.

pathway in yeast resulted in an increased proliferation of ER membrane, which was hypothesized to compensate for the loss of SRP-dependent targeting of proteins to the ER membrane [20]. Unlike other eukaryotic

cells, yeast contain both a well-characterized posttranslational system and two cotranslational systems for targeting proteins to the ER membrane which may explain why proliferation of the ER is sufficient to prevent the yeast from dying. However, downregulation in HeLa cells of two SRP polypeptides essential for protein targeting did not result in a lethal phenotype suggesting that there might be as yet uncharacterized pathways to target proteins to the ER membrane in all eukaryotic cells [21].

The question of mRNA partitioning and the interchange of ribosomes between the free and bound forms may help resolve the confusion generated by experiments showing that the SRP pathway is not essential. It is well known that in cells there are two pools of ribosomes: those translating cytoplasmic proteins reside primarily in the cytoplasm, and those translating secretory and integral membrane proteins partitioned predominantly on the ER membrane. Once the mRNA is fully translated and the nascent protein has been transferred through or integrated into the ER membrane, the bound ribosomes can be released from the membrane and join the pool of free ribosomes in the cell. However, biochemical data suggest that membrane-bound ribosomes are not easily released from the membrane even after release of the mRNA and newly synthesized protein. In the past, the observation that mRNAs encoding abundant cytoplasmic proteins were present in membrane fractions was usually assumed to result from contamination. However, experiments involving cDNA microarray screens suggest that membrane fractions contain an unexpectedly large number of mRNAs encoding cytosolic proteins [22]. This result raises the possibility that mRNA partitioning is more complex than initially postulated (Figure 8.2).

To address the possible functions of ER-bound ribosomes, Potter and Nicchitta incubated rough microsomal membranes charged with posttermination ribosomes in a ribosome-free reticulocyte lysate and demonstrated the *de novo* synthesis of both secreted and cytoplasmic proteins. Surprisingly, they demonstrated that synthesis of a protein that did not contain a signal sequence led to the release of these RNCs from the membrane [14]. This experiment revealed another ribosome-partitioning pathway now denoted the elongation-coupled ribosome release (E-CRR) pathway (Figure 8.2). Unlike the signal sequence-mediated SRP pathway, which is a positive regulator for partitioning of ribosomes carrying secretory mRNAs to the ER membrane, the E-CRR pathway involves a negative regulator that releases ribosomes translating cytoplasmic proteins from the ER. The operation of the two processes together better describes the distribution of the two pools of mRNAs, especially the observation that ribosomes charged with cytosolic mRNAs can be found on the ER. What it does not address is the mechanism by which mRNAs encoding cytosolic proteins are initially targeted to the ER-bound membranes.

Perhaps as originally speculated in the 1960s [3], information for partitioning can be encoded in the mRNA [23]. By acting in concert, SRP, E-CRR, and mRNA targeting may result in very high fidelity and efficient separation of proteins destined for the ER from ones that should remain in the cytoplasm (Figure 8.2). The use or reuse of membrane-bound ribosomes is one method of efficiently targeting signal peptide-containing polypeptides with substoichiometric amounts of SRP. It might also permit mRNAs that encode secreted proteins to be translated many times (and each time by several membrane-bound ribosomes) without needing to be targeted by SRP. However, the use of membrane-bound ribosomes might also result in translation of an mRNA encoding a cytoplasmic protein which would be problematic if translocation via membrane-bound ribosomes is SRP independent. Thus, E-CRR may permit optimal use of membrane-bound ribosomes while minimizing mistargeting of newly synthesized proteins. Furthermore, a postsynthesis quality control step that disposed of mistargeted proteins would further increase the ultimate overall fidelity of the targeting reaction [24]. Regardless, it is clear that there are still unresolved issues regarding the mechanisms used to target proteins to the ER.

## IV. Targeting of Proteins to the ER Is Regulated by Unusual GTPases

Due primarily to experiments using cell-free systems, the details of the molecular mechanisms of the SRP pathway are beginning to be elucidated. SRP and both the α- and β-subunits of the SR contain well-defined GTPase domains. Contrary to most well-studied GTP-binding proteins, SRP and SRα alone in solution have low affinities for nucleotides, and the nucleotides that occupy the active sites can be readily exchanged [25]. However, the situation changes when SRP binds to SRα by a "concerted switch" mechanism in which both components require a bound GTP molecule [25]. Structures solved for the complex between the bacterial homologues of the GTP-binding domains of SRP and SRα demonstrated that the guanine nucleotide bound in one active site provided an integral hydrogen bond to the reciprocal GTP-binding site and vice versa [26, 27], resulting in a very high-affinity interaction. Once SRP mediates high-affinity binding to the signal sequence, the regulation of the GTP binding and hydrolysis activity of SRP and SRα is the major determinant for the fidelity of the SRP-targeting pathway [28]. Multiple steps in the pathway alter the GTP-binding equilibrium such that nascent chain substrates bound to SRP are delivered to the translocation pore in the ER membrane. For instance, ribosome binding by SRP increases the affinity of SRP54 for GTP [29]. Binding of

the RNC-bound SRP to SRα further stimulates the binding of GTP by SRP54 [30]. Interestingly, even though GTP-dependent SRα binding was originally found to promote the dissociation of SRP from the ribosome [31], it was later discovered that in the absence of detergent, SRP remained bound to both the RNC and SRα until the Sec61 translocation complex was present [32]. This requirement for proximity to the Sec61 complex is presumed to ensure that the RNC is delivered to a translocation-competent location. Finally, GTP hydrolysis by SRP and/or SRα is required for the dissociation of SRP from SRα [33].

The SRP54–SRα interaction has a half-life of >6 h *in vitro*. Thus, complex formation makes this step of the targeting process effectively unidirectional [25]. The GTPase-driven interaction between SRα-SRP54 also serves as the initiating step for protein translocation, governing both the targeting of the RNC and the subsequent formation of the ribosome-translocon junction [34]. Although recognition by SRP and binding to the SR results in the unidirectional transfer of the RNC to the translocon, the nascent protein must still be subsequently recognized as a translocation substrate or the process can be aborted by the E-CRR pathway described earlier.

There is at present no adequate explanation for SRα being a peripheral membrane protein in eukaryotes and most prokaryotes. In fact, in many prokaryotes the protein appears to be a cytoplasmic protein. Therefore, to target an SRP-bound RNC to the membrane SRα must interact with an integral membrane protein. To date this protein, SRβ, has been unambiguously identified only in eukaryotes.

### A. Characterization of the SRβ Subunit of the SR

Sequencing of the SRβ cDNA revealed two important features. First, a 19-amino acid stretch of hydrophobic residues located at the N-terminus of the protein, which was proposed and later confirmed through carbonate extractions and proteolysis topology experiments, to act as a Type I signal-anchor sequence [34, 35]. Second, the remainder of SRβ is a GTPase domain located on the cytoplasmic side of the ER. The SRβ GTPase is distantly related to those of SRP54 and SRα although SRβ has a nanomolar affinity for GTP. Only the GTP-bound form of canine SRβ binds to SRα via the SRX domain of SRα [36] suggesting that GTPase activity also regulates the binding of SRβ to SRα. Elucidation of the structure of the yeast SRX bound to a mutant of SRβ lacking the transmembrane segment confirmed that the heterodimerization of SR is highly dependent on the SRβ GTP/GDP switch, with dimerization occurring only when SRβ is GTP-bound [37]. Although it binds GTP very tightly, canine SRβ has no detectable GTPase activity alone or bound to SRα [36]. It is possible therefore

that SRβ requires a GTPase activating protein (GAP) and/or a guanine nucleotide exchange factor (GEF) for proper GTPase function.

At the ER membrane, SRβ can interact with a number of components to regulate transfer of the RNC to the translocation complex but how GTP binding and hydrolysis regulates this transfer is still unknown. The cellular GTP concentration is ~0.5 mM and therefore, it is assumed that SRβ is in the GTP-bound state on the ER membrane [36, 37]. One hypothesis for the GTP regulation of SRβ proposes that the ribosome is the GAP for SRβ [38]. Another possible GAP for SRβ is Sec61β, the β-subunit of the Sec61 complex. In this scenario (Figure 8.3, Model 1), the ribosome of the SRP-bound RNC would make initial contact prior to binding with a component of the Sec61 translocation complex. The function of SR would be to scan the membrane and interact with an SRP-bound RNC on the ER membrane to functionally engage it with a translocon. The interaction with the translocon would trigger both the release of the signal sequence from SRP54 and GTP hydrolysis by SRP54 and SRα [38]. After transfer of the nascent chain to the translocon, interaction of SRβ with the translocon via Sec61β (or a change in the ribosome activating an inherent GAP activity) would trigger hydrolysis of GTP by SRβ releasing SRα from the membrane [37].

In an alternative model (Figure 8.3, Model 2), it has been suggested that initial targeting of the SRP-bound RNC to the SR occurs via an interaction with SRα. In this scenario, once the SRP-bound RNC docks to SRα, GAP activity provided by the SRX domain of SRα dissociates SRβ from SRα to transfer the complex of SRα, SRP and the RNC to the translocation complex [39–41]. Subsequent hydrolysis of GTP by SRα would release SRα from SRP allowing it to rebind to SRβ. The main difference between the two models is the timing and function of hydrolysis of GTP by SRβ. In model 1, it is used to release targeting factors from the RNC after the tight seal has formed between the ribosome and the translocon, whereas in model 2, it releases SRα from SRβ to transfer the RNC to the translocon.

The data obtained from eukaryotic cell-free systems and in yeast provide a promising start toward the full elucidation of SRβ function in SRP-dependent cotranslational translocation. Unfortunately, little information can be obtained from the prokaryotic system as there are no known SRβ homologues in prokaryotes. The SRα homologue in prokaryotes (FtsY) shows conservation with the eukaryotic protein only in the GTPase domain. Moreover, the detailed mechanism of targeting of FtsY to the membrane is still elusive. In most Gram-negative bacteria, FtsY contains an acidic N-terminal region that can mediate the targeting of FtsY to membranes via an interaction with a phospholipid in the membrane and an unidentified protein [42]. However, most Gram-positive bacteria contain an FtsY protein that is composed of the highly conserved GTPase domain without

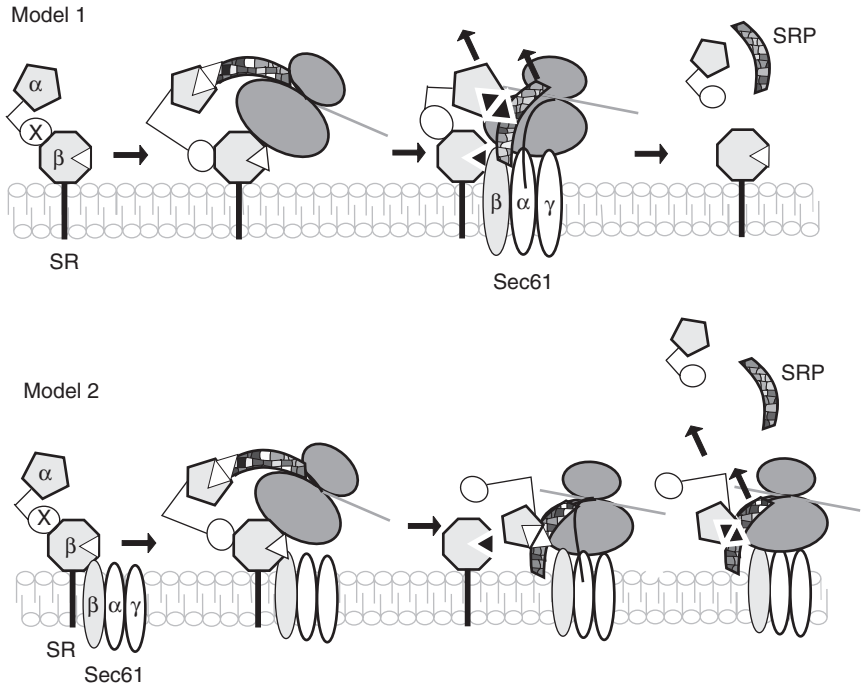



FIG. 8.3. Two models to describe the possible function of the SRβ GTPase in the SRP pathway. In both models, SRα binds SRβ when SRβ is bound to GTP (white triangle). Interaction of the signal sequence with SRP54 increases the affinity of SRP54 for GTP and binds the RNC to the membrane via the tight GTP-mediated interaction between SRP54-GTP and SRα-GTP. Model 1: Once at the membrane, the SR-bound RNC migrates to a vacant Sec61 translocon. The nascent polypeptide interacts with Sec61α and, after the tight seal with the translocon forms, the nascent polypeptide is transferred to the translocon. Once the RNC is bound to the translocon, Sec61β acts as a GAP for SRβ resulting in GTP hydrolysis (GDP, black triangles) concomitant with hydrolysis of GTP by both SRα and SRP54 dissociating all of the targeting proteins from the ribosome [37]. Model 2: SR is docked at Sec61 complexes prior to the arrival of the SRP bound RNC. Once the SRP-bound RNC is bound to SRα, the SRX domain acts as a GAP for SRβ [39]. Hydrolysis of GTP by SRβ releases SRα, allowing the SRP-bound RNC to bind tightly to the translocon. Transfer of the nascent polypeptide to the translocon stimulates hydrolysis of GTP by SRα and SRP54 releasing them from each other and the RNC. In both models the low affinity of SRα and SRP54 for GDP results in spontaneous exchange for GTP via a transient but stable empty state. Finally, an unidentified GEF (not shown) exchanges GDP for GTP in SRβ recreating the binding site for SRα.

the acidic N-terminal region. These homologues of FtsY are mostly cytoplasmic proteins but are presumed to interact with the bacterial inner membrane transiently through an unknown mechanism. A small subset of bacteria has evolved a version of FtsY that contains an authentic

transmembrane segment, thereby combining features of the eukaryotic SRα-SRβ heterodimer into one protein. The main difference is that this transmembrane version of FtsY contains a single GTPase domain. Thus, although the prokaryotic SR has proven to be as complex in mechanism as that of its eukaryotic counterpart, no prokaryotic system employs three GTPases.

### B. Docking the Ribosome on the Translocon

The last step in targeting the RNC to the ER involves attaching the ribosome to the translocation machinery in the ER. This docking reaction appears to trigger GTP hydrolysis by SRP and SRα, resulting in the transfer of the signal peptide to the translocon where it initially interacts with Sec61α [43]. This appears to be the last proofreading step in signal peptide recognition and therefore the last step at which targeting can be aborted. Once the ribosome functionally engages the translocon a tight seal is formed that prevents the passage of ions between the ER lumen and the cytoplasm during polypeptide translocation and commits the nascent protein to transport across or into the ER membrane. The protein(s) involved in establishing the ion impermeable connection between the ribosome and the membrane have not been identified. Structural analysis of the translocon components that remain bound to the ribosome after detergent solubilization of the membrane [44, 45] suggests that the protein(s) involved are lost during isolation of the complex as the nascent chain appears accessible from the cytoplasmic side of the membrane. In contrast, fluorescence studies demonstrate that in intact membranes the nascent chain can only be accessed from the lumenal side of the ER membrane and then only after a gate in the translocon has opened [46]. The reconstitution of a minimally functioning translocation complex required the Sec61 heterotrimer (Sec61α, Sec61β, and Sec61γ) and SR, although the translocation of specific nascent polypeptides required additional proteins such as TRAM [47]. Unfortunately fluorescence studies on translocons reconstituted into membranes from purified components that would likely reveal that the tight seal is not present have not been performed. Nevertheless, it appears that the last step in the targeting process is not completely reconstituted in these systems.

In order to downregulate the host immune response, cytomegalovirus has evolved a system that, for selected substrates, defeats commitment to translocation subsequent to functional engagement of the ribosome with the translocon. However, this system appears to return the nascent protein to the cytoplasm only after translation has completed and is coupled to

destruction of the target protein [48]. Thus, this system is technically not part of the protein targeting step.

After translation and translocation has completed for a given polypeptide, the now empty (i.e., nontranslating) ribosome remains bound with high affinity to the Sec61 translocation complex, as described earlier. When the ER membrane was fully saturated with nontranslating ribosomes, newly added translating ribosomes bound to these membranes via Sec61 complexes but binding required SRP and SR [49]. Thus, it appears that SRP and SR can direct the binding of RNCs to a pool of Sec61 translocation complexes that are otherwise not available for high-affinity binding to ribosomes. It is possible that trimers (or tetramers) of Sec61 heterotrimers constitute the high-affinity binding sites for ribosomes but an equilibrium with lower affinity monomeric/dimeric Sec61 heterotrimers ensures the availability of SRP/SR-dependent binding sites for RNCs [49].

## V. Structure–Function Analysis

With the major components of the SRP pathway identified and the molecular functions ascribed to the individual components, the examination of the SRP pathway has moved toward identifying the precise molecular mechanisms involved. Structural analyses of the SRP and SR$\alpha$ GTPases revealed the molecular details of the novel regulatory mechanism described above. However, unlike the GTPases described above that could also be studied as enzymes [35, 41], most of the proteins involved in targeting RNCs to the ER are not enzymes. Structural biology is, therefore, integral to the determination of molecular mechanism(s) involved in targeting proteins to membranes and during the last decade, many structures have been solved for interacting components of the SRP cycle. However, work in this area is far from complete and the functional implications of many of the structures that have been solved remain elusive.

### A. Signal Recognition Particle

The SRP is a ubiquitous ribonucleoprotein having been found in the cells of all types of organisms tested, including mammalian [50], plant [51], yeast [52], bacterial [53], and archaebacterial cells [54]. The evolutionary conservation of the SRP highlights its importance in cellular function. Comparison of these SRPs reveals the essential core constituents as the GTPase (SRP54) that binds the signal sequence as described earlier, and an RNA (SRP RNA). SRP obtained from canine microsomes was the first SRP studied and represents the SRP archetype even though it is larger and

contains more protein components than the SRPs found in many other organisms. Canine SRP is composed of a 300-nucleotide SRP RNA, and six polypeptides classified by their molecular masses: SRP9, SRP14, SRP19, SRP54, SRP68, and SRP72 [50]. Early studies showed that the particle was rod shaped and composed of three domains [55, 56]. Based on this structure, a model was proposed in which SRP spanned from the nascent chain exit site on the ribosome to the site of protein synthesis where it could slow translation by directly interfering with peptide bond formation or tRNA binding. Cryo-EM studies of the SRP bound to the ribosome largely confirm this overall structure, domain organization and ribosome-binding site for SRP. However, the end of SRP that interferes with translation was shown to bind to a region of the ribosome known to be the binding site for elongation factors [57]. Therefore, it is now believed that canine SRP arrests translation by directly competing with elongation factors for ribosome binding [57, 58]. Higher resolution information about the structure of SRP and its receptor has come from the analysis of individual components but, as yet, it has not proven possible to integrate all of the structural information into a comprehensive high-resolution structure for SRP from any organism.

## B. SRP RNA

Initially, eukaryotic SRP was believed to be a complex of six different polypeptides (a heterohexamer); however, purified canine SRP consistently generated a high ratio of absorbance at 260 nm compared to the absorbance at 280 nm. This characteristic of SRP led to the discovery of an integral RNA molecule, found to be a critical component for the activity of eukaryotic SRP [59]. The SRP RNA, as detected by electron spectroscopic imaging, runs along the entire length of the rod-like canine SRP [55] supporting the hypothesis that one major function of the SRP RNA is to act as a scaffold to combine the various functions of the SRP protein components into a single complex. The eukaryotic SRP RNA has a well conserved, highly base-paired secondary structure (Figure 8.4; reviewed in [60]). SRP RNA can be structurally (and functionally, see later) divided into two major domains; the Alu domain, composed of double-helical regions H2, H3, and H4; and, the S domain, composed of H6, H7, H8, and part of H5 [61]. The Alu and S domains are linked by the long H5 double-helical region, which is believed to have flexible regions that can act as hinges to modulate the ternary structure of SRP at different stages of the SRP pathway. Region H8 is present in every known SRP RNA, including those of archaeal and prokaryotic origins, some of which are composed of only this region [60] therefore, H8 represents a minimal SRP RNA. Of the eukaryotic SRP RNAs, *Saccharomyces cerevisiae* and other yeasts have a

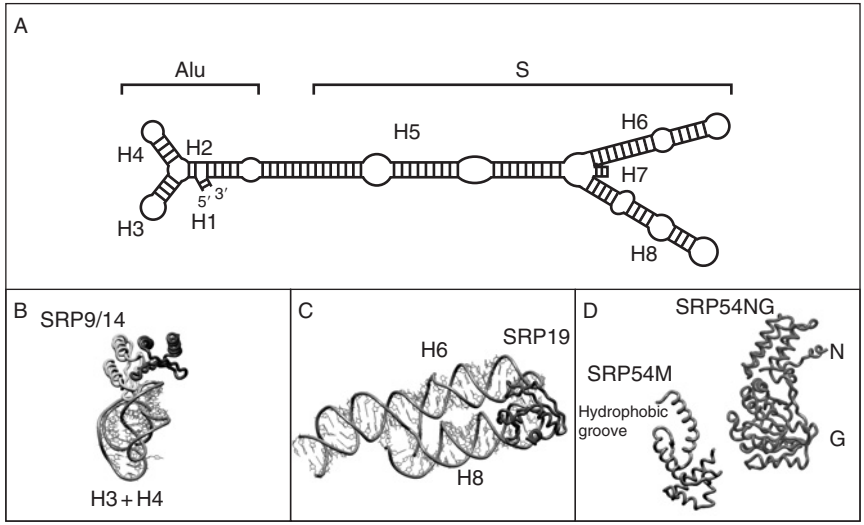



FIG. 8.4. Structures of the eukaryotic SRP components. (A) A schematic representing the conserved regions of secondary structure of SRP RNA. The mammalian SRP RNA was used as the prototypic model for eukaryotic SRP RNA [61]. Highly base-paired, double-helical regions are labeled H1 to H8. The Alu and S fragments are indicated by square brackets. The number and placement of the unpaired, loops, or bulges is not accurate. (B) The structure solved for the cocrystal of the heterodimer formed by human SRP9 (light) and human SRP14 (dark) bound to a region of the Alu fragment of SRP RNA (including H3 and H4). (PDB: 1E8O [75]). (C) The ternary complex of SRP19 bound to the distal ends of H6 and H8. (PDB: 1L9A [84]). (D) The only crystal structure solved for eukaryotic SRP54 is the M domain from *Homo sapiens* (SRP54M, PDB: 1QB2 [99]). The structure solved for the NG domain of *Sulfolobus solfataricus* SRP54 is shown for reference (SRP54NG, PDB: 1QZX [104]). Ribbon diagrams of protein and RNA structures were generated using Insight II (Accelrys Software, Inc., San Diego, United States).

more complex RNA with several additional double-helical segments extending from H5. The function of these extra helical regions is not known.

Limited nuclease digestion of canine SRP resulted in the separation of two distinct ribonucleoprotein (RNP) complexes: one containing the Alu domain of the RNA bound to the SRP9 and SRP14 polypeptides, and the other containing the S domain bound to SRP19, SRP54, SRP68, and SRP72 [62]. Enzymatic and chemical probing assays indicated that the S domain could be further subdivided; SRP19 and SRP54 localized to the ends of H6 and H8, and SRP68 and SRP72 bound to H5 and the adjacent regions of H6 and H8 [63, 64]. These results are consistent with the SRP RNA acting as a scaffold. However, many bacteria have an SRP that contains only one polypeptide suggesting that the SRP RNA must have an additional function in order for it to have been retained as a necessary component of all SRPs.

The SRP RNA could mediate the binding between SRP and the ribosome. The structure of canine SRP bound to a RNC solved by cryo-electron microscopy (cryo-EM) indicated contacts between SRP and the ribosome that were solely mediated by the SRP RNA [57]. Also, it has been demonstrated that, in an *in vitro* translation reaction, SRP RNA increased the affinity of SRP54 for ribosome-bound nascent chains consistent with a role for SRP RNA in binding ribosomes [65].

Another potential role for the SRP RNA is that it could alter or stabilize the tertiary structure of the SRP polypeptides. In *E. coli*, it was found that at least one domain of the SRP54 homologue protein had a more cooperative thermal denaturation profile in the presence of the H8 region of its cognate SRP RNA [66], suggesting that the structure of this protein domain was more tightly folded in the presence of the SRP RNA. Also, SRP72 was found to bind to SRP68 in an SRP RNA-dependent manner [67] indicating that the SRP RNA facilitates the exposure or changes the conformation of the SRP68–SRP72 dimerization interface.

The accessibility of the SRP RNA to chemical modifying reagents was altered at specific stages of the SRP pathway [64] suggesting that the SRP RNA may contribute to regulating the function of SRP by inducing conformational changes in the particle. Additionally, the SRP RNA has been proposed to participate directly in the binding and/or recognition of the signal sequence based on a structure solved for a cocrystal of a fragment of *E. coli* SRP RNA bound to a single domain of an SRP protein that indicated that H8 was in proximity to the signal sequence binding site [68]. The *bona fide* functions of SRP RNA must be elucidated before they can be incorporated into a more comprehensive molecular mechanism for SRP function.

### C. SRP9 AND SRP14

The SRP9/14 heterodimer is responsible for arresting translation elongation of ribosome-bound, signal sequence-containing nascent chains [69, 70]. The SRP9 and SRP14 polypeptides must bind to each other and form a stable heterodimer [71], referred to as SRP9/14, before they can bind with high affinity to the Alu domain of SRP RNA [72, 73]. Even though they share only limited sequence homology, SRP9 and SRP14 are structurally similar [74]. They are both composed of two α-helices and three β-strands, which form a concave six strand antiparallel β-sheet in the heterodimer, with the α-helices arranged on the convex side [74]. The concave face of SRP9/14 straddles the compact, helical stacks formed by H3 and H4 of the Alu domain of SRP RNA [75]. The structure solved for SRP9/14 provided new insight into the mode of RNA binding by this heterodimer

(Figure 8.4B), but a full structural explanation of the role of SRP9/14 in SRP-dependent protein targeting is still lacking. The RNA-bound SRP9/14 structure did eliminate the possibility that SRP9/14 organizes the Alu domain such that it mimics a tRNA molecule to facilitate elongation arrest.

The SRP9/14 heterodimer appears to be exclusive to higher eukaryotes although there is evidence of potential homologues in yeast. While *S. cerevisiae* lack an SRP9 homologue, they do have a homologue of SRP14 (56% sequence similarity to canine SRP14) [76], and it has been shown that the yeast homologue of SRP14 can homodimerize and bind to a region of the yeast SRP RNA corresponding to the Alu domain [77]. Moreover, the homodimer of the yeast SRP14 homologue facilitated an arrest in nascent chain elongation, both *in vitro* and *in vivo* [78].

### D. SRP68 AND SRP72

The SRP68 and SRP72 polypeptides form an SRP68/72 heterodimer [71] although, SRP68 binding to the SRP RNA is a prerequisite for the heterodimerization of these two polypeptides [67]. RNA binding by SRP68 is believed to tether SRP72 to the SRP RNA through protein–protein contacts. Reconstituted canine SRP lacking the SRP68/72 heterodimer was unable to promote the translocation of SRP-dependent polypeptides [70]. Consistent with this finding, when SRP68/72 was modified with N-ethylmaleimide (NEM), which alkylates free sulfhydryl groups of cysteine residues, the resulting reconstituted SRP was also defective in polypeptide translocation [79]. However, SRP containing the NEM-modified SRP68/72 was still competent for binding to SR. Cryo-EM analysis of ribosome-bound SRP located the SRP68/72 in a putative hinge region of H5 of SRP RNA [57]. It was proposed that SRP68/72 either stabilizes or modulates the conformation of this region of the SRP complex to transmit information of the signal sequence binding state of SRP to the SRP9/14 region of SRP to promote elongation arrest. Interestingly, reconstituted SRP lacking SRP68/72 was also deficient in elongation arrest activity [70], which could be consistent with a role for SRP68/72 in the positioning of the Alu domain of SRP on the ribosome. Alternatively, since SRP that lacks SRP68/72 is deficient in so many activities perhaps the primary role of SRP68/72 is to directly influence the relative positioning of SRP9/14 and the other functional components of SRP at the correct locations on the ribosome.

SRP68 was found to contain three regions of primary sequence that are similar to those conserved in guanine nucleotide dissociation stimulators [80]. This may indicate a role for the heterodimer, or SRP68 on its own, in the regulation of one or more of the GTPases present in the SRP pathway.

The structure of the SRP68/72 heterodimer remains unsolved and the volume allocated for the SRP68/72 heterodimer in the SRP-ribosome cryo-EM structure appears insufficient to adequately incorporate a complex that accounts for 60% of the protein mass of eukaryotic SRP. Further structural and functional work is required to clarify the role of this relatively large component of the eukaryotic SRP.

## E. SRP19

The SRP19 polypeptide binds to the distal ends of H6 and H8 of the SRP RNA [81, 82]. It is a small protein composed of a three-strand antiparallel β-sheet packed against two α-helices [83]. A ternary structure solved for the archaeal homologue of SRP19 bound to H6 and H8 of SRP RNA demonstrated that SRP19 brings these two double-helical regions of the SRP RNA together, allowing them to bind to each other (Figure 8.4C) [84, 85]. The nuclease protection patterns of SRP19 on SRP RNA and of SRP19 with SRP54 on SRP RNA were identical [63] so it was originally proposed that SRP19 tethered SRP54 to the SRP RNA via protein–protein interactions, analogous to the SRP68/72 heterodimer. However, it was later found that SRP54 could bind isolated H8 of the *E. coli* SRP RNA in the absence of SRP19 [86]. The cocrystal of the RNA-binding domain of the *E. coli* homologue of SRP54 bound to the H8 region of its cognate SRP RNA indicated that the binding sites of SRP19 and SRP54 were distinct [68]. The addition of SRP19 to SRP RNA increased the affinity of SRP54 for SRP RNA 100-fold [87]. Together this data suggests that SRP19 causes a stabilization of a binding site for SRP54 on H8 either by the direct binding of SRP19 to H8 or by bringing H8 into proximity of H6.

Homologues of SRP19 have only been found in organisms that have an SRP RNA that contains both H6 and H8, that is primarily eukaryotes and archaea. It is not clear what additional functions/activities the H6 region and SRP19 provide for these SRP molecules but they must not be critical to the core functions of SRP as it was lost during the evolution of many prokaryotic species.

## F. SRP54

The SRP54 subunit is the only SRP component that has been found in all SRP homologues highlighting the importance of SRP54 in the SRP pathway. SRP54 directly mediates signal sequence recognition and nascent chain-binding activities of the SRP complex [88–90]. In fact, it has been demonstrated that even without SRP RNA, isolated SRP54 discriminated and bound to signal sequences [91]. The primary sequence of SRP54 has

been arbitrarily divided into the N (N-terminal), G (GTPase), and M (methionine-rich) regions. Much of the structural and biochemical research into SRP54 has been on the bacterial homologues (described in Chapter 1), therefore only the key points will be reiterated here with a specific focus on eukaryote-specific findings.

The N region is a bundle of four antiparallel α-helices [92], which is believed to regulate or respond to the occupancy of the GTP-binding site of the G region. Mutations in a conserved seven residue motif, which comprises a major link between the N region and the G region (via side chain interactions), caused defects in signal sequence binding [93] and might indicate a possible role for the N region in contributing to the SRP54-signal sequence binding activity. However, reconstituted SRP which lacked the N and G regions of SRP54 was still observed to bind the nascent chain and promote elongation arrest [65, 94, 95]. Therefore, the N region may have an indirect effect on the apparent affinity for signal sequences by altering the interaction of SRP54 with the translating ribosome.

The structural studies performed on the N and G regions of bacterial SRP54 have indicated that, although the N and G regions are structurally distinct, they are organized as a single domain, referred to as the NG domain (Figure 8.4D) [92]. Within this domain the G region structure is typical of other low molecular weight GTPases. The SR has a highly homologous NG domain and it is through the respective NG domains that SRP54 and SR heterodimerize [31, 65]. The structure solved for the heterodimer of the bacterial NG domains from the SRP54 and SR homologues indicated that extensive regions of both the N and G regions (including the bound nucleotide) contributed to the heterodimerization interface [26, 27].

There is evidence that the NG domain may be participating in the signal sequence binding activity of SRP54. Alkylation of cysteine residues of the NG domain in intact SRP by NEM treatment blocked nascent chain-binding [94]. Also, SRP reconstituted with a modified SRP54 lacking the NG domain had a lower affinity for signal sequences compared to complete SRP [65]. In *E. coli*, an isolated NG domain was found to cross-link to a peptide carrying a signal sequence [96], although the exact site of signal sequence binding in the NG domain is still unknown. It is possible that the role of the eukaryotic NG domain in signal sequence binding is direct, or it may stabilize the M domain.

The NG domain has also been observed to enhance the binding of SRP54 to the SRP RNA [87]. In *E. coli*, isolated NG domain was observed to bind the *E. coli* SRP RNA by nitrocellulose filtration [97] although the affinity of this reaction was 20-fold worse than the affinity of the full-length SRP54. Clearly, if RNA binding by NG is relevant the domain is providing only an accessory binding site.

The M region is a 22-kDa C-terminal domain of SRP54 that was originally named for the relatively high methionine content in the mammalian protein [98]. Most of the SRP54 homologues from other organisms also have an unusually high occurrence of methionine residues in the M regions although several examples exist, especially in thermophilic organisms, where the distribution of methionine residues is not as striking. A structure (accounting for 120 of the 180 total amino acids) solved for the M region from human SRP54 revealed a hydrophobic groove lined by α-helices (Figure 8.4D) [99]. In the crystal used to solve this structure, the hydrophobic groove of one M molecule was occupied by an α-helix from a neighboring M molecule. It was hypothesized that this extra α-helix could be mimicking a signal sequence and that the mechanism of binding in the hydrophobic groove was a good model for the signal sequence binding activity of the M region. Originally, the accepted theory for the signal sequence binding activity was that the methionine side chains of the M region acted as ''bristles'' protruding from amphipathic helices that could accommodate a wide variety of sequences that may be found in signal sequences (reviewed in [100]). This theory was proposed based on limited homology between the M region and calmodulin, a protein which does utilize methionine residues to bind to peptides. The methionine residues of the human M region do contribute to the hydrophobic groove but do not localize to structural elements that would support a specific role of the methionine residues in nascent chain binding. Therefore, it appears that it is the shape and hydrophobic character of the groove that provides the functionality to bind a wide range of signal sequences [99].

The hydrophobic groove of the human M region is deep and elongated, widened at each end with a slight constriction in the middle [99]. The dimensions of this groove would be compatible with binding a signal sequence in an α-helical conformation and would cause a slight kink in the nascent chain at the constriction site [99]. In the structure of the bacterial M region from *Thermus aquaticus*, a similar hydrophobic groove was observed although it had a more open conformation but, similar to the human protein described above, it was occupied by a hydrophobic extended (rather than helical) loop from a crystallographically adjacent M region [101]. If the human M region has the structural plasticity to adopt a similar more open conformation then it may also bind to a hydrophobic sequence that is not a strict α-helix. Also, the 60 amino acids not included in the crystallographic study of the structure of the human M region may provide additional structural information about the nascent chain-binding activity.

The M region of SRP54 is also the major binding site for H8 of SRP RNA [94], and it was observed that isolated M regions contained all of the necessary structural elements to bind H8 [95]. Interestingly, it was found

that the addition of *E. coli* SRP RNA to its cognate M region reversed a global destabilization of the M region in response to signal peptide addition [66]. This would suggest that the M region undergoes a conformational change on signal sequence binding that is stabilized through the binding of SRP RNA to the M region.

Scanning transmission electron microscopy (STEM) was used to determine the three-dimensional structure of intact canine SRP54 by angular reconstitution using 200 STEM images of single molecules [102]. This structure showed two unequal masses connected by a linker. The larger mass, likely the NG domain, was wedged-shaped and the smaller mass, the M region, was crescent shaped. The proximity of the M region with the NG domain and the long linker between the two masses led to the proposal that the M region could move relative to NG to open/close in a ''clamp''-like fashion to bind nascent chains. This could explain the inhibition of nascent chain binding to SRP54 by NEM-mediated alkylation of the NG domain [94]. Consistent with a model of M and NG being in proximity, addition of a peptide bearing a signal sequence to SRP54 caused a stabilization of the nucleotide-free state of the GTP-binding site of the NG domain [30].

### G. SRP21, an SRP Subunit Unique to Yeast

The initial purification of SRP from *S. cerevisiae* identified a novel 21-kDa protein, SRP21 [76], that appears to be a subunit of active SRP [78]. Deletion of the gene encoding this subunit resulted in a decrease in SRP production indicating that SRP21 is required for the stable expression of *S. cerevisiae* SRP [76]. SRP21 was found to be required for the assembly and nuclear export of yeast SRP [103] which may be an indirect result of SRP21 affecting the stability of yeast SRP rather than it functioning as a yeast specific SRP export factor. This subunit may provide a function that complements the extra double-helical regions of *S. cerevisiae* SRP RNA (which is approximately twice the size of most other eukaryotic SRP RNAs). However, homologues of SRP21 have been found in other fungal organisms, including fungi that have an SRP RNA molecule of the ''normal'' eukaryote size (250–270 nucleotides). Thus, despite the advantages of yeast as a model organism less is known about SRP21 than for the other SRP proteins.

### H. Engaging the Translocation Complex

The final step of the SRP protein targeting pathway is the docking of the RNC to a competent Sec61 translocation complex. As discussed earlier, the exact mechanism of the transfer of the translating ribosome from the

SRP–SR complex to the translocation complex is still unknown; however, some structural data have elucidated some intriguing aspects of this process. The structure of the SRP-bound RNC, as solved by cryo-EM, indicated that SRP54 was positioned immediately adjacent to the peptide exit site of the large subunit of the ribosome, overlapping a region of the ribosome that is the binding site for the translocation complex [57]. When a soluble version of the SR heterodimer was added and the structure of this new complex was solved by cryo-EM, a delocalization of the NG domain of SRP54 in response to SR binding was observed [58]. This SR-dependent conformational change in SRP54 opens up the binding site for the translocation complex on the ribosome surface which could promote the functional attachment and transfer of the RNC to the translocation complex in order that, as translation of the nascent polypeptide continues, the polypeptide moves into or across the ER membrane. Many of the details of this final step remain in contention, including determination of the complete complement of components necessary for a fully functional translocation complex and the gating mechanism that regulates access to and from the translocation pore (Chapters 10–12).

## VI. Conclusions

It is clear that the mechanisms used to target nascent proteins to the ER are complex and tightly regulated. However, they are also ancient, with analogous systems found in every domain of life. The degree of specialization in eukaryotes including the involvement of three unusual GTPases and a complex multiprotein SRP suggests that the targeting system may also be an important site of regulation of gene expression and of other cellular functions associated with membrane protein biogenesis. Furthermore, there is significant evidence for alternative pathways and specialized regulatory systems for specific translocation substrates. Thus, although many of the core functions and components of the targeting system have been identified and the molecular mechanisms are being elucidated by a combination of biochemical and structural approaches, there is still much to understand about how protein targeting to the ER is regulated in eukaryotes.

### Acknowledgments

The authors' research is supported by Canadian Institute of Health Research (CIHR) grant FRN-10490 to DWA and studentship from National Science and Engineering Research Council (F.V.). DWA holds the Canada Research Chair in Membrane Biogenesis.

## References


1. Kim, P.K., Janiak-Spens, F., Trimble, W.S., Leber, B., and Andrews, D.W. (1997). Evidence for multiple mechanisms for membrane binding and integration via carboxyl-terminal insertion sequences. *Biochemistry* 36:8873–8882.
2. Redman, C.M., and Sabatini, D.D. (1966). Vectorial discharge of peptides released by puromycin from attached ribosomes. *Proc Natl Acad Sci USA* 56:608–615.
3. Adelman, M.R., Sabatini, D.D., and Blobel, G. (1973). Ribosome-membrane interaction. Nondistructive disassembly of rat liver rough microsomes into ribosomal and membranous components. *J Cell Biol* 56:206–229.
4. Lingappa, V.R., Katz, F.N., Lodish, H.F., and Blobel, G. (1978). A signal sequence for the insertion of a transmembrane glycoprotein. *J Biol Chem* 253:8667–8670.
5. von Heijne, G. (1985). Signal sequences. The limits of variation. *J Mol Biol* 184:99–105.
6. Sabatini, D.D., Blobel, G., Nonomura, Y., and Adelman, M.R. (1971). Ribosome-membrane interaction: structural aspects and functional implications. *Adv Cytopharmacol* 1:119–129.
7. Blobel, G., and Dobberstein, B. (1975). Transfer of Proteins across Membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol* 67:835–851.
8. Lingappa, V.R., Chaidez, J., Yost, S., and Hedgpeth, J. (1984). Determinants for protein localization: β-lactamase signal sequence directs globin across microsomal membranes. *Proc Natl Acad Sci USA* 81:456–460.
9. Andrews, D.W., Perara, E.P., Lesser, C., and Lingappa, V.R. (1988). Sequences beyond the cleavage site influence signal peptide function. *J Biol Chem* 263:15791–15798.
10. Martoglio, B., and Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. *Trends Cell Biol* 8:410–415.
11. Perara, E., and Lingappa, V.R. (1985). A former amino terminal signal sequence engineered to an internal location directs translocation of both flanking protein domains. *J Cell Biol* 101:2292–2301.
12. Mize, N.K., Andrews, D.W., and Lingappa, V.R. (1986). A stop transfer sequence recognizes receptors for nascent chain translocation across the endoplasmic reticulum membrane. *Cell* 47:711–719.
13. Wessels, H.P., and Spiess, M. (1988). Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. *Cell* 55:61–70.
14. Potter, M.D., and Nicchitta, C.V. (2000). Regulation of ribosome detachment from the mammalian endoplasmic reticulum membrane. *J Biol Chem* 275:33828–33835.
15. Lingappa, V.R. (1991). More than just a channel: provocative new features of protein traffic across the ER membrane. *Cell* 65:527–530.
16. Andrews, D.W., and Johnson, A.E. (1996). The translocon: more than a hole in the ER membrane? *TIBS* 21:365–369.
17. Walter, P., Ibrahimi, I., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum I. Signal recognition protein (SRP) binds to *in-vitro*-assembled polysomes synthesizing secretory protein. *J Cell Biol* 91:545–550.
18. Siegel, V., and Walter, P. (1988). The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. *EMBO J* 7:1769–1775.
19. Sato, R., Yang, J., Wang, X., Evans, M.J., Ho, Y.K., Goldstein, J.L., and Brown, M.S. (1994). Assignment of the membrane attachment, DNA binding, and transcriptional activation domains of sterol regulatory element-binding protein-1 (SREBP-1). *J Biol Chem* 269:17267–17273.

20. Mutka, S.C., and Walter, P. (2001). Multifaceted physiological response allows yeast to adapt to the loss of the signal recognition particle-dependent protein-targeting pathway. *Mol Biol Cell* 12:577–588.
21. Ren, Y.G., Wagner, K.W., Knee, D.A., Aza-Blank, P., Nasoff, M., and Deveraux, Q.L. (2004). Differential regulation of the TRAIL death receptors DR4 and DR5 by the signal recognition particle. *Mol Biol Cell* 15:5064–5074.
22. Kopczynski, C.C., Noordermeer, J.N., Serano, T.L., Chen, W.Y., Pendleton, J.D., Lewis, S., Goodman, C.S., and Rubin, G. (1998). A high throughput screen to identify secreted and transmembrane proteins involved in *Drosophila* embryogenesis. *Proc Natl Acad Sci USA* 95:9973–9978.
23. Nicchitta, C.V., Lerner, R.S., Stephens, S.B., Dodd, R.D., and Pyhtila, B. (2005). Pathways for compartmentalizing protein synthesis in eukaryotic cells: the template-partitioning model. *Biochem Cell Biol* 83:687–695.
24. Simon, K., Perara, E., and Lingappa, V.R. (1987). Translocation of globin fusion proteins across the endoplasmic reticulum membrane in *Xenopus laevis* oocytes. *J Cell Biol* 104:1165–1172.
25. Rapiejko, P.J., and Gilmore, R. (1997). Empty site forms of the SRP54 and SR alpha GTPases mediate targeting of ribosome-nascent chain complexes to the endoplasmic reticulum. *Cell* 89:703–713.
26. Egea, P.F., Shan, S.O., Napetschnig, J., Savage, D.F., Walter, P., and Stroud, R.M. (2004). Substrate twinning activates the signal recognition particle and its receptor. *Nature* 427:215–221.
27. Focia, P.J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004). Heterodimeric GTPase core of the SRP targeting complex. *Science* 303:373–377.
28. Millman, J.S., and Andrews, D.W. (1997). Switching the model: a concerted mechanism for GTPases in protein targeting. *Cell* 89:673–676.
29. Bacher, G., Lutcke, H., Jungnickel, B., Rapoport, T.A., and Dobberstein, B. (1996). Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature* 381:248–251.
30. Miller, J.D., Wilhelm, H., Gierasch, L., Gilmore, R., and Walter, P. (1993). GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature* 366:351–354.
31. Connolly, T., and Gilmore, R. (1989). The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide. *Cell* 57:599–610.
32. Song, W., Raden, D., Mandon, E.C., and Gilmore, R. (2000). Role of Sec61alpha in the regulated transfer of the ribosome-nascent chain complex from the signal recognition particle to the translocation channel. *Cell* 100:333–343.
33. Connolly, T., Rapiejko, P.J., and Gilmore, R. (1991). Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. *Science* 252:1171–1173.
34. Miller, J.D., Tajima, S., Lauffer, L., and Walter, P. (1995). The β subunit of the signal recognition particle receptor is a transmembrane GTPase that anchors the α subunit, a peripheral membrane GTPase, to the endoplasmic reticulum membrane. *J Cell Biol* 128:273–282.
35. Legate, K.R., Falcone, D., and Andrews, D.W. (2000). Nucleotide-dependent binding of the GTPase domain of the signal recognition particle receptor beta-subunit to the alpha-subunit. *J Biol Chem* 275:27439–27446.
36. Schwartz, T., and Blobel, G. (2003). Structural basis for the function of the β subunit of the eukaryotic signal recognition particle receptor. *Cell* 112:793–803.

37. Helmers, J., Schmidt, D., Glavy, J.S., Blobel, G., and Schwartz, T. (2003). The β-subunit of the protein-conducting channel of the endoplasmic reticulum functions as the guanine nucleotide exchange factor for the β-subunit of the signal recognition particle receptor. *J Biol Chem* 278:23686–23690.
38. Bacher, G., Pool, M., and Dobberstein, B. (1999). The ribosome regulates the GTPase of the β-subunit of the signal recognition particle receptor. *J Cell Biol* 146:723–730.
39. Schlenker, O., Hendricks, A., Sinning, I., and Wild, K. (2006). The structure of the mammalian signal recognition particle (SRP) receptor as prototype for the interaction of small GTPases with Login domains. *J Biol Chem* 281:8898–8906.
40. Legate, K.R., and Andrews, D.W. (2001). Assembly strategies and GTPase regulation of the eukaryotic and *Escherichia coli* translocons. *Biochem Cell Biol* 79:593–601.
41. Legate, K.R., and Andrews, D.W. (2003). The β-subunit of the signal recognition particle receptor is a novel GTP-binding protein without intrinsic GTPase activity. *J Biol Chem* 278:27712–27720.
42. Millman, J.S., Qi, H.Y., Vulcu, F., Bernstein, H.D., and Andrews, D.W. (2001). FtsY binds to the *Escherichia coli* inner membrane via interactions with phosphatidylethanolamine and membrane proteins. *J Biol Chem* 276:25982–25989.
43. McCormick, P.J., Miao, Y., Shao, Y., Lin, J., and Johnson, A.E. (2003). Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. *Mol Cell* 12:329–341.
44. Menetret, J.-F., Hegde, R.S., Heinrich, S.U., Chandramouli, P., Ludtke, S.J., Rapoport, T.A., and Akey, C.W. (2005). Architecture of the ribosome-channel complex derived from native membranes. *J Mol Biol* 348:445–457.
45. Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., and Akey, C.W. (2000). The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 6:1219–1232.
46. Hamman, B.D., Hendershot, L.M., and Johnson, A.E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the luminal end of the translocon pore before and early in translocation. *Cell* 92:747–758.
47. Gorlich, D., and Rapoport, T.A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75:615–630.
48. Lilley, B.N., and Ploegh, H.L. (2004). A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 429:834–840.
49. Schaletzky, J., and Rapoport, T.A. (2006). Ribosome binding to and dissociation from translocation sites of the ER membrane. *Mol Biol Cell* 17:3860–3869.
50. Walter, P., and Blobel, G. (1980). Purification of a membrane associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc Natl Acad Sci USA* 77:7112–7116.
51. Campos, N., Palau, J., Torrent, M., and Ludevid, D. (1988). Signal recognition-like particles are present in maize. *J Biol Chem* 263:9646–9650.
52. Hann, B.C., Poritz, M.A., and Walter, P. (1989). *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* contain a homologue to the 54-kD subunit of the signal recognition particle that in *S. cerevisiae* is essential for growth. *J Cell Biol* 109:3223–3230.
53. Romisch, K., Webb, J., Herz, J., Prehn, S., Frank, R., Vingron, M., and Dobberstein, B. (1989). Homology of 54K protein of signal-recognition particle, docking protein and two *E. coli* proteins with putative GTP-binding domains. *Nature* 340:478–482.
54. Moll, R., Schmidtke, S., and Schafer, G. (1999). Domain structure, GTP-hydrolyzing activity and 7S RNA binding of *Acidianus ambivalens* Ffh-homologous protein suggest an SRP-like complex in archaea. *Eur J Biochem* 259:441–448.

55. Andrews, D.W., Walter, P., and Ottensmeyer, F.P. (1987). Evidence for an extended 7SL RNA structure in the signal recognition particle. *EMBO J* 6:3471–3477.
56. Andrews, D.W., Walter, P., and Ottensmeyer, F.P. (1985). Structure of the signal recognition particle by electron microscopy. *Proc Natl Acad Sci USA* 82:785–789.
57. Halic, M., Becker, T., Pool, M.R., Spahn, C.M., Grassucci, R.A., Frank, J., and Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* 427:808–814.
58. Halic, M., Gartmann, M., Schlenker, O., Mielke, T., Pool, M.R., Sinning, I., and Beckmann, R. (2006). Signal recognition particle receptor exposes the ribosomal translocon binding site. *Science* 312:745–747.
59. Walter, P., and Blobel, G. (1982). Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* 299:691–698.
60. Poritz, M.A., Strub, K., and Walter, P. (1988). Human SRP RNA and *E. coli* 4.5S RNA contain a highly homologous structural domain. *Cell* 55:4–6.
61. Larsen, N., and Zwieb, C. (1991). SRP-RNA sequence alignment and secondary structure. *Nucleic Acids Res* 19:209–215.
62. Gundelfinger, E.D., Krause, E., Melli, M., and Dobberstein, B. (1983). The organization of the 7SL RNA in the signal recognition particle. *Nucleic Acids Res* 11:7363–7374.
63. Siegel, V., and Walter, P. (1988). Binding sites of the 19-kDa and 68/72-kDa signal recognition particle (SRP) proteins on SRP RNA as determined in protein-RNA ''footprinting.'' *Proc Natl Acad Sci USA* 85:1801–1805.
64. Andreazzoli, M., and Gerbi, S.A. (1991). Changes in 7SL RNA conformation during the signal recognition particle cycle. *EMBO J* 10:767–777.
65. Zopf, D., Bernstein, H.D., and Walter, P. (1993). GTPase domain of the 54-kD subunit of the mammalian signal recognition particle is required for protein translocation but not for signal sequence binding. *J Cell Biol* 120:1113–1121.
66. Zheng, N., and Gierasch, L.M. (1997). Domain interactions in *E. coli* SRP: stabilization of M domain by RNA is required for effective signal sequence modulation of NG domain. *Mol Cell* 1:79–87.
67. Lutcke, H., Prehn, S., Ashford, A.J., Remus, M., Frank, R., and Dobberstein, B. (1993). Assembly of the 68- and 72-kD proteins of signal recognition particle with 7S RNA. *J Cell Biol* 121:977–985.
68. Batey, R.T., Rambo, R.P., Lucast, L., Rha, B., and Doudna, J.A. (2000). Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* 287:1232–1239.
69. Siegel, V., and Walter, P. (1986). Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature* 320:81–84.
70. Siegel, V., and Walter, P. (1985). Elongation arrest is not a prerequisite for secretory protein translocation across the microsomal membrane. *J Cell Biol* 100:1913–1921.
71. Walter, P., and Blobel, G. (1983). Disassembly and reconstitution of signal recognition particle. *Cell* 34:525–533.
72. Janiak, F., Walter, P., and Johnson, A.E. (1992). Fluorescence-detected assembly of the signal recognition particle: binding of the two SRP protein heterodimers to SRP RNA is noncooperative. *Biochemistry* 31:5830–5840.
73. Strub, K., and Walter, P. (1990). Assembly of the Alu domain of the signal recognition particle (SRP): dimerization of the two protein components is required for efficient binding to SRP RNA. *Mol Cell Biol* 10:777–784.
74. Birse, D.E., Kapp, U., Strub, K., Cusack, S., and Aberg, A. (1997). The crystal structure of the signal recognition particle Alu RNA binding heterodimer, SRP9/14. *EMBO J* 16:3757–3766.

75. Weichenrieder, O., Wild, K., Strub, K., and Cusack, S. (2000). Structure and assembly of the Alu domain of the mammalian signal recognition particle. *Nature* 408:167–173.
76. Brown, J.D., Hann, B.C., Medzihradszky, K.F., Niwa, M., Burlingame, A.L., and Walter, P. (1994). Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. *EMBO J* 13:4390–4400.
77. Zwieb, C., van Nues, R.W., Rosenblad, M.A., Brown, J.D., and Samuelsson, T. (2005). A nomenclature for all signal recognition particle RNAs. *RNA* 11:7–13.
78. Mason, N., Ciufo, L.F., and Brown, J.D. (2000). Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J* 19:4164–4174.
79. Siegel, V., and Walter, P. (1988). Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* 52:39–49.
80. Althoff, S., Selinger, D., and Wise, J.A. (1994). Molecular evolution of SRP cycle components: functional implications. *Nucleic Acids Res* 22:1933–1947.
81. Zwieb, C. (1991). Interaction of protein SRP19 with signal recognition particle RNA lacking individual RNA-helices. *Nucleic Acids Res* 19:2955–2960.
82. Watanabe, M., and Blobel, G. (1989). SecB functions as a cytosolic signal recognition factor for protein export in *E. coli*. *Cell* 58:695–705.
83. Wild, K., Sinning, I., and Cusack, S. (2001). Crystal structure of an early protein-RNA assembly complex of the signal recognition particle. *Science* 294:598–601.
84. Oubridge, C., Kuglstatter, A., Jovine, L., and Nagai, K. (2002). Crystal structure of SRP19 in complex with the S domain of SRP RNA and its implication for the assembly of the signal recognition particle. *Mol Cell* 9:1251–1261.
85. Hainzl, T., Huang, S., and Sauer-Eriksson, A.E. (2002). Structure of the SRP19 RNA complex and implications for signal recognition particle assembly. *Nature* 417:767–771.
86. Poritz, M.A., Bernstein, H.D., Strub, K., Zopf, D., Wilhelm, H., and Walter, P. (1990). An *E. coli* ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. *Science* 250:1111–1117.
87. Diener, J.L., and Wilson, C. (2000). Role of SRP19 in assembly of the Archaeoglobus fulgidus signal recognition particle. *Biochemistry* 39:12862–12874.
88. Kurzchalia, T.V., Wiedmann, M., Girshovich, A.S., Bochkareva, E.S., Bielka, H., and Rapoport, T.A. (1986). The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle. *Nature* 320:634–636.
89. Krieg, U.C., Walter, P., and Johnson, A.E. (1986). Photocrosslinking of the signal sequence of nascent preprolactin to the 54-kilodalton polypeptide of the signal recognition particle. *Proc Natl Acad Sci USA* 83:8604–8608.
90. Wiedmann, M., Kurzchalia, T.V., Bielka, H., and Rapoport, T.A. (1987). Direct probing of the interaction between the signal sequence of nascent preprolactin and the signal recognition particle by specific cross-linking. *J Cell Biol* 104:201–208.
91. Hauser, S., Bacher, G., Dobberstein, B., and Lutcke, H. (1995). A complex of the signal sequence binding protein and the SRP RNA promotes translocation of nascent proteins. *EMBO J* 14:5485–5493.
92. Freymann, D.M., Keenan, R.J., Stroud, R.M., and Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature* 385:361–364.
93. Newitt, J.A., and Bernstein, H.D. (1997). The N-domain of the signal recognition particle 54-kDa subunit promotes efficient signal sequence binding. *Eur J Biochem* 245:720–729.
94. Lutcke, H., High, S., Romisch, K., Ashford, A.J., and Dobberstein, B. (1992). The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO J* 11:1543–1551.

95. High, S., and Dobberstein, B. (1991). The signal sequence interacts with the methionine-rich domain of the 54-kD protein of signal recognition particle. *J Cell Biol* 113:229–233.
96. Cleverley, R.M., and Gierasch, L.M. (2002). Mapping the signal sequence-binding site on SRP reveals a significant role for the NG domain. *J Biol Chem* 277:46763–46768.
97. Buskiewicz, I., Kubarenko, A., Peske, F., Rodnina, M.V., and Wintermeyer, W. (2005). Domain rearrangement of SRP protein Ffh upon binding 4. 5S RNA and the SRP receptor FtsY. *RNA* 11:947–957.
98. Bernstein, H.D., Poritz, M.A., Strub, K., Hoben, P.J., Brenner, S., and Walter, P. (1989). Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature* 340:482–486.
99. Clemons, W.M., Jr., Gowda, K., Black, S.D., Zwieb, C., and Ramakrishnan, V. (1999). Crystal structure of the conserved subdomain of human protein SRP54M at 2.1 A resolution: Evidence for the mechanism of signal peptide binding. *J Mol Biol* 292:697–705.
100. Zheng, N., and Gierasch, L.M. (1996). Signal sequences: the same yet different. *Cell* 86:849–852.
101. Keenan, R.J., Freymann, D.M., Walter, P., and Stroud, R.M. (1998). Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell* 94:181–191.
102. Czarnota, G.J., Andrews, D.W., Farrow, N.A., and Ottensmeyer, F.P. (1994). A structure for the signal sequence binding protein SRP54: 3D reconstruction from STEM images of single molecules. *J Struct Biol* 113:35–46.
103. Grosshans, H., Deinert, K., Hurt, E., and Simos, G. (2001). Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export. *J Cell Biol* 153:745–762.
104. Rosendal, K.R., Wild, K., Montoya, G., and Sinning, I. (2003). Crystal structure of the complete core of archaeal signal recognition particle and implications for interdomain communication. *Proc Natl Acad Sci USA* 100:14701–14706.

# 9

# *The Translocation Apparatus of the Endoplasmic Reticulum*

MARTINA ALKEN • RAMANUJAN S. HEGDE

*Cell Biology and Metabolism Branch*
*National Institute of Child Health and Human Development*
*National Institutes of Health*
*Bethesda, MD 20892, USA*

## I. Abstract

Eukaryotic proteins destined for the cell surface, extracellular space, or compartments of the secretory pathway are first translocated across or inserted into the endoplasmic reticulum (ER) membrane at sites termed translocons [1–4]. The essential feature of ER translocons is a protein-conducting channel formed by the highly conserved Sec61 complex. Together with several accessory components, the Sec61 complex recognizes translocation substrates, provides a gated conduit for transport across the membrane, and regulates access to the lipid bilayer for membrane protein integration. These combined activities endow translocons with the remarkable capacity to direct the proper biogenesis and topology for a tremendously diverse set of secretory and membrane protein substrates. How is this complex feat accomplished? In this chapter, we subdivide the protein translocation process into a series of decisive mechanistic steps taken by a substrate during its transit across or insertion into the membrane. The translocon components implicated in each step and their proposed mechanisms of action are considered with an eye toward particularly important gaps in our understanding of protein translocation into the ER.



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25009-7

## II. Translocons Receive Substrates via Two Distinct Pathways

The first step in translocation is the targeting of a substrate to the translocon. Depending on the substrate and the organism, targeting is achieved in two qualitatively different ways (Figure 9.1). In the mammalian system, almost all secretory and membrane proteins are recognized and targeted to the membrane cotranslationally (i.e., while they are being synthesized by cytosolic ribosomes). By contrast, a substantial proportion of proteins in yeast can be targeted posttranslationally, after their complete synthesis and release into the cytosol.

In the cotranslational targeting pathway (see Chapter 8), substrates are recognized when the first hydrophobic domain, either a signal sequence or a

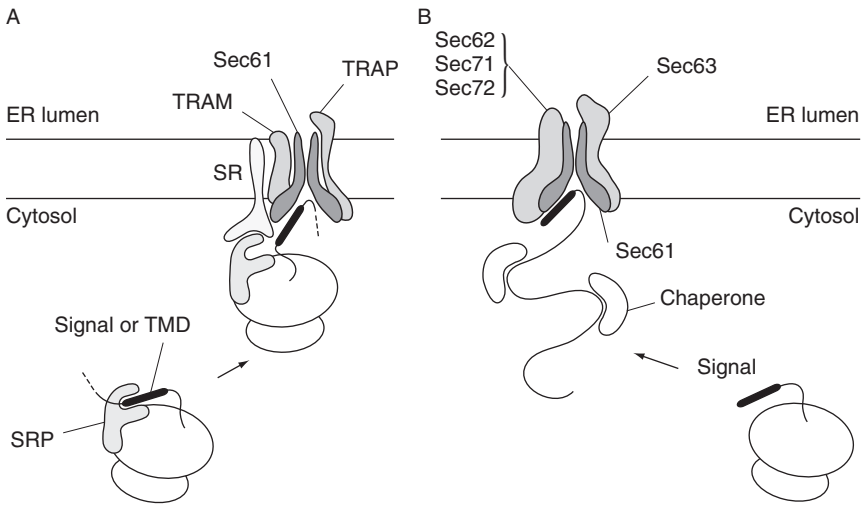



FIG. 9.1. Two modes of substrate delivery to translocons at the ER. (A) Cotranslational targeting. The SRP binds the hydrophobic regions of N-terminal signal sequences and TMDs as they emerge from the ribosome (bottom). The RNC-SRP complex is delivered to a translocon at the ER (top) via a GTP-dependent interaction between SRP and its receptor (SR). The principal components of the mammalian cotranslational translocon are indicated, with the highly conserved Sec61 complex forming the central protein-conducting channel. (B) Posttranslational targeting. Polypeptides with N-terminal signal sequences of lower hydrophobicity fail to be recognized efficiently by SRP in some organisms (such as *Saccharomyces cerevisiae*). In this case, the nascent chain is bound by various cytosolic chaperones that keep it in a loosely folded configuration until signal sequence recognition by an ER translocon initiates chaperone disengagement and protein translocation. The principal components of the posttranslational translocon from *S. cerevisiae* are indicated, with the Sec61 complex again forming the central protein-conducting channel.

transmembrane domain (TMD), emerges from the translating ribosome. The signal recognition particle (SRP) binds to the exposed hydrophobic domain, slows translation via further contacts with the ribosome, and targets the entire ribosome-nascent chain (RNC) complex to the endoplasmic reticulum (ER)-localized SRP receptor [5]. The RNC is subsequently transferred to the translocon (in a very poorly understood step), ensuring the delivery of nascent chains for translocation at an early stage in substrate synthesis.

In yeast, SRP fails to efficiently recognize signal sequences whose hydrophobicity falls below a certain threshold [6, 7]. These polypeptides are therefore not targeted to the ER cotranslationally, but instead are bound by cytosolic chaperones, including Hsp70, its cofactor Ydj1p, and the Tric/CCT chaperonin complex [8–10]. These factors maintain signal-containing proteins in an unfolded flexible conformation prior to their transfer to the ER translocon. The mechanisms that coordinate chaperone release with delivery of substrates to the translocon remain poorly understood, but may be dependent on a functional signal sequence. Thus, in both pathways, a signal- or TMD-containing polypeptide is presented to an unengaged ER translocon in a configuration (either unfolded or while still being synthesized) that facilitates subsequent translocation through a channel of limited size.

## III. Substrate Recognition by the ER Translocon Is a Decisive Step in Protein Translocation

On delivery to the translocon, substrates are recognized and discriminated a second time using the same element (either a signal sequence or TMD) initially used for targeting. This recognition step serves at least three purposes. First, it allows secretory and membrane proteins to be discriminated from nonsecretory proteins that may have inadvertently been targeted to a translocon and therefore serves as a ''proofreading'' step to improve fidelity of sorting [11, 12]. Second, the engagement of a translocon by a signal or TMD presumably prepares the translocon for protein translocation by gating the channel from a closed to an open configuration [13]. And finally, the orientation in which the signal or TMD is recognized determines the topology achieved by the substrate [14, 15]. Thus, the translocon must not only recognize and interact productively with a signal or TMD but must also orient this domain with respect to the lipid bilayer.

The orientation taken by the signal or TMD is a critical step in the final outcome of the polypeptide substrate (Figure 9.2; see also Figure 9.3A and B). An N-terminal cleavable signal sequence is generally thought to

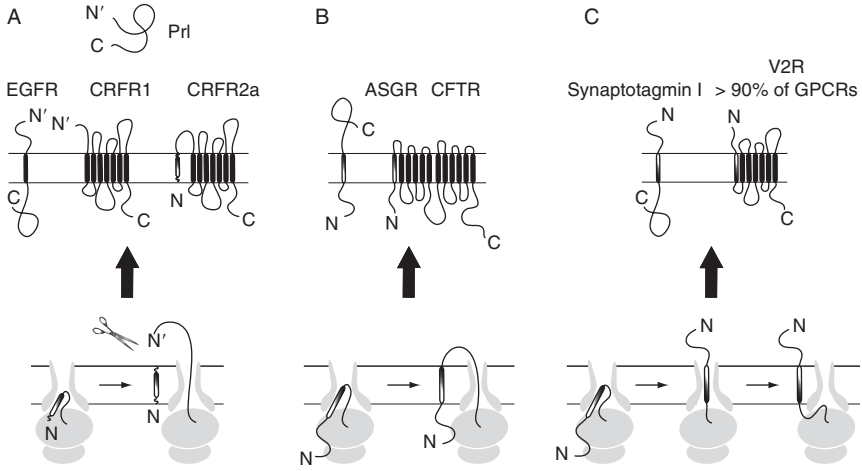



FIG. 9.2. The initiation step of protein translocation influences protein topology. (A) An N-terminal signal sequence is generally recognized by the translocon in a ''looped'' orientation with the N-terminus facing the cytosol (lower-left diagram). On cleavage of the signal sequence by signal peptidase (indicated by the scissors), the new N-terminus of the mature polypeptide (N') is committed to translocation into the ER lumen. Hence, nearly all proteins with an N-terminal signal sequence have the N-terminus of the mature protein in a noncytosolic location. Examples include secretory proteins like the hormone prolactin, simple membrane proteins like the EGF-receptor (EGFR), and complex membrane proteins like the G-protein–coupled receptor for the corticotropin releasing factor (CRFR). Although rare, some proteins (like the 2a isoform of the CRFR) may retain the N-terminal signal sequence in their final structure [134]. (B and C) Proteins whose targeting is achieved by a TMD are made into membrane proteins. Depending on features of the TMD and its flanking sequences, the orientation it acquires in the translocon is either with the N-terminus facing the cytosol (panel B) or translocated into the lumen (panel C). When the targeting TMD acquire the topology depicted in panel B, it is often known as a ''signal anchor'' or type I signal anchor sequence. Panel C depicts a ''reverse signal anchor'' or type II signal anchor. Examples of membrane proteins that utilize these targeting mechanisms are shown. Note that unlike the N-terminal signal sequence, the targeting element is part of the final protein structure.

be positioned with the N-terminus facing the cytosol and the C-terminal region oriented toward the ER lumen (Figure 9.2A). This orientation exposes the signal cleavage site to the lumenal side of the translocon (where the active site of signal peptidase is located [16]) and positions the nascent polypeptide in a ''looped'' configuration. When positioned this way, the N-terminus of the mature polypeptide can access the ER lumen and translocate across the membrane through the aqueous protein-conducting channel in the translocon (Figure 9.3A); failure to achieve this looped configuration results in a lack of

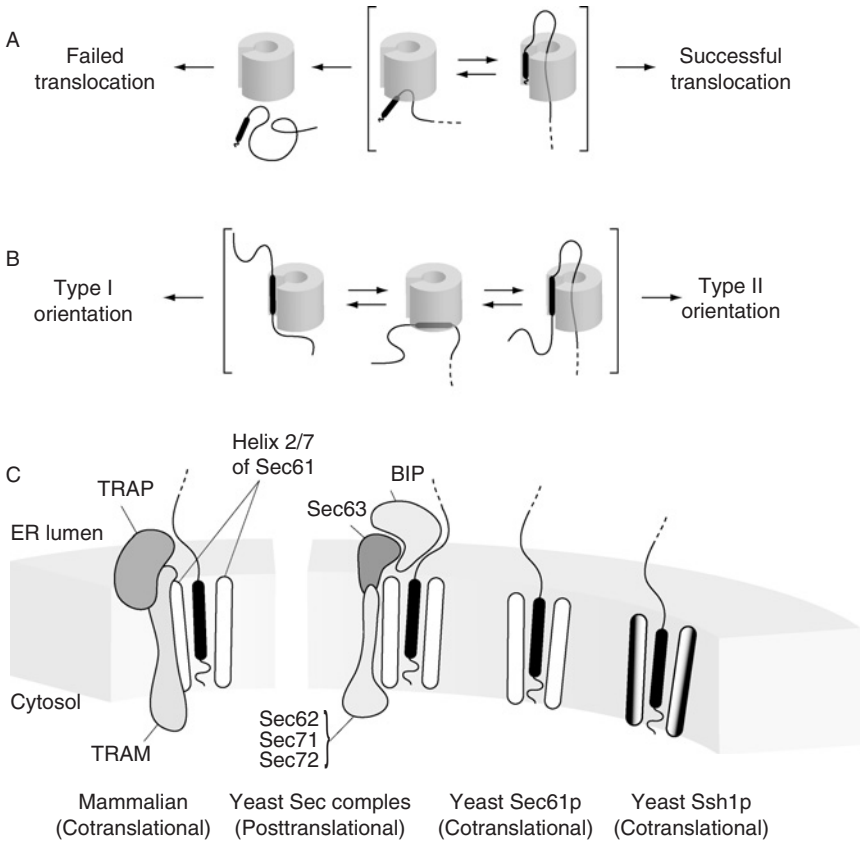



FIG. 9.3. The signal sequence-translocon interaction. In each of the diagrams, the lumen is on top, and the cytosol on the bottom. For simplicity, the ribosome (which would be bound to the bottom of each translocon) is not shown in any of these pictures. (A) Shown in the brackets is the initial interaction between a signal sequence and a translocon (depicted as a cylinder). This interaction is thought to be dynamic (indicated by the double arrows), and occurs during ongoing protein translation. If excessive protein is synthesized and folding initiates on the cytosolic side before a ''looped'' orientation is achieved (left diagram), the polypeptide fails to be translocated. Otherwise, the polypeptide can commit to being successfully translocated. Signal sequences from different proteins seem to achieve the looped orientation and hence successfully gate the translocon with differing efficiencies [135]. The membrane bilayer is not shown for simplicity. (B) As with the signal sequence in panel A, the interaction between a TMD and the translocon is thought to be similarly dynamic [136]. The possible configurations that may be sampled by a TMD are shown in the brackets, with the arrows indicating an ability to interconvert among the different states. Depending on which configuration is stabilized sufficiently long to permit lateral movement into the lipid bilayer, the substrate can be committed to either final orientation. The membrane bilayer is not shown for simplicity. (C) The signal-translocon in different systems is shown in molecular detail. In each case, the view is essentially from ''within'' the channel of the translocon, looking out laterally toward the

translocation. Thus, a cleavable N-terminal signal sequence is employed by a wide range of secretory and membrane proteins whose N-terminus needs to be translocated to the noncytosolic side of the membrane (Figure 9.2A).

A TMD can be positioned in one of two orientations after its recognition by the translocon (Figure 9.2B and C). If the N-terminal side of the TMD faces the cytosol, it is often referred to as a signal anchor (or sometimes a type II signal anchor). This orientation is analogous to a cleaved signal sequence, and the nascent chain similarly acquires a ''looped'' configuration at this early step in translocation (Figure 9.2B). If the TMD is oriented with the N-terminal side facing the lumen, it is called a reverse (or sometimes a type I) signal anchor (Figure 9.2C). In this case, the N-terminal domain to the TMD must be translocated to the lumenal side of the membrane to permit such an orientation, while the C-terminal domain to the TMD must remain in the cytosol. Regardless of the orientation, the TMD is eventually integrated into the membrane (i.e., moved from the proteinaceous environment of the translocon to the lipid bilayer). Both types of signal anchors are employed widely in the biogenesis of both single- and multispanning membrane proteins, examples of which are indicated in Figure 9.2B and C.

From this discussion, it should be apparent that the initial recognition by a translocon of the targeting element (either a signal sequence or TMD) directly determines the portion of a nascent polypeptide that is translocated across the membrane versus retained on the cytosolic side (i.e., the *topology* of the final polypeptide). In the posttranslational translocation pathway, this signal-translocon interaction is also likely to be the sole discriminatory event in segregating secretory from nonsecretory proteins [17, 18]. Even in cotranslational translocation, signal recognition by the translocon may be the decisive discriminatory event since SRP-mediated targeting may not be obligatory for translocation of substrates whose mRNAs are already

lipid bilayer. Thus, the only portion of the translocation channel that is displayed is the putative lateral exit site formed by helices 2 and 7 of Sec61α (white bars). In the mammalian cotranslational system, the signal sequence (black bar) is positioned in the helix 2/7 interface such that portions flanking its hydrophobic region may contact portions of TRAM and the TRAP complex for additional stabilization. The analogous interaction in the yeast posttranslational system also involves the helix 2/7 interface of Sec61p. Here, additional stabilizing interactions are provided by Sec62p, Sec63p, Sec71p, Sec72p, and/or the lumenal chaperone BiP (also called Kar2p). Cotranslational translocation in yeast may rely solely on the helix 2/7 interaction since homologues of TRAM and TRAP do not appear to exist. This may explain why only a subset of signal sequences (presumably those that do not need stabilizing factors) are directed into the cotranslational pathway in yeast [6, 7]. Furthermore, it could also explain why a second cotranslational translocon formed by a Sec61 homologue (Ssh1p), which has a somewhat different helix 2/7 site (indicated by shaded bars), displays different signal sequence specificity [62].

docked on ER-bound polysomes [11]. Consistent with this notion, yeast mutants depleted of SRP are initially defective in translocation and grow slowly, but substantially recover their translocation capacity and growth rate over time [19, 20]. A plausible explanation for this observation is that cotranslational translocation substrates are initially very slow to engage the translocon in the absence of SRP, but become much more efficient after the first (presumably random, and hence rare) targeting event delivers the mRNA to the ER membrane. Indeed, *in vitro* experiments have shown that a polypeptide can be successfully translocated in a signal-discriminatory manner in the complete absence of SRP, provided that the RNC complex is first docked at the translocon [11, 21–24]. Thus, recognition of a hydrophobic-targeting sequence by the ER translocon is both an obligatory and decisive step in the initiation of all modes of protein translocation.

## IV. The Remarkable Diversity of Sequences Recognized by the Translocon

Even though essentially every substrate translocated across (or inserted into) the ER engages the translocon, the domains that are recognized (signal sequences and TMDs) share no sequence motifs or homology whatsoever [25–27]. Not only are signal sequences distinctly different from TMDs, but each motif is itself highly variable. N-terminal signal sequences are usually ~15–45 amino acids in length and are often considered to have a three-domain structure: a nonhydrophobic and often basic n-region, followed by a central hydrophobic core of ~8–12 residues (h-region), and ending in a c-region that often contains helix-breaking and small uncharged residues.

TMDs are generally longer than signal sequences and have a hydrophobic membrane-spanning domain of at least ~16 residues. This hydrophobic domain is often flanked by charged residues, the asymmetric distribution of which often correlates with its final orientation in the membrane (with basic residues favoring the cytosolic side, the so-called ''positive-in'' rule [28, 29]. Because the TMDs used for targeting need not be at the beginning of the protein, the N-terminal domain to the TMD can be of highly variable size (Figure 9.2B and C). Thus, the only feature that is considered common to all signal sequences and TMDs is a ~8-residue hydrophobic region uninterrupted by charges. For this reason, the hydrophobic region of the signal and TMD is generally considered the principal feature that is recognized by the translocon and engages it for subsequent events in substrate translocation.

What then is the functional role of sequence diversity in signals and TMDs? In the case of TMDs, at least two functions are clear. First, statistical

and mutagenesis studies have demonstrated that the length, hydrophobicity, and flanking charge distribution of the TMD all influence its orientation in the translocon and hence final protein topology [14, 15, 30]. Second, any structural and functional roles of the TMD in the final membrane protein inevitably constrain the sequences that are allowable. In the case of signal sequences, role(s) for sequence variability among substrates remains poorly understood. Only relatively recently have functional differences among signal sequences begun to emerge from a collection of seemingly disparate studies from multiple experimental systems [31, 32].

The reason for this brief discussion on the sequence features of TMDs and signals is that it highlights several important questions regarding translocon function that are considered in subsequent sections. First, how are the rather vague *shared* features of signals and TMDs (primarily hydrophobicity) recognized by the translocon despite their enormous sequence diversity? Second, how are the *variable* regions of signals and TMDs interpreted by the translocon to impart substrate-specific features such as orientation of a TMD? And how might this sequence diversity be physiologically exploited by the cell for differential regulation of translocation in a substrate-specific manner? In considering these and other questions, it is therefore important to keep in mind that even though most of our current knowledge comes from the study of a very few model examples, translocons are in fact designed to handle remarkable substrate diversity.

## V. The Machinery of Signal Sequence Recognition

What does a signal ''see'' when it interacts with the translocon? The answer to this question turns out to depend on the organism, substrate, and mode of translocation. This is because the translocation apparatus is not exactly the same in all organisms, and multiple translocons of differing compositions are likely to exist even within a single cell [1–4]. Nonetheless, it is generally thought that due to their similar hydrophobic character, signal sequences and TMDs of either orientation are all recognized in approximately the same way and at the same general site in a translocon. Furthermore, the high degree of evolutionary conservation in the essential targeting and translocon machinery (i.e., the SRP system and the Sec61 complex) suggests that the fundamental steps in translocation (such as signal sequence recognition) will vary little across experimental systems [33, 34]. With this in mind, let us now consider what is known about the machinery of signal and TMD recognition.

Two types of substrates have been analyzed over the years. In the mammalian system, RNCs of defined length and substrate composition

are assembled with pancreatic ER-derived microsomes. In the yeast system, a posttranslational substrate is stalled in its translocation across yeast-derived microsomes. Insight into signal sequence and TMD recognition by the translocon in each system comes primarily from fluorescence, cross-linking, and reconstitution studies. Although no single approach or system provides a complete view, a composite can be deduced by combining the resulting data with a reasonable assumption of high evolutionary conservation of basic mechanisms.

In the fluorescence approach, nascent chains containing a fluorescent amino acid incorporated in the signal sequence were used to probe the local environment at different stages in translocation [35]. These pioneering studies demonstrated that during cotranslational translocation, the signal sequence of a model secretory protein is in an aqueous environment continuously during its transit through the ribosomal tunnel and engagement of the ER translocation [13, 36]. On translocon engagement, the probe in the signal, although still in an aqueous environment, was no longer accessible to ions from the cytosolic side of the membrane [13, 36]. It was presumed that the signal must therefore occupy a nonhydrophobic space at the membrane shielded from the cytosol, consistent with it interacting directly with the translocon.

Coincident with these studies, cross-linking approaches had identified several integral membrane proteins in proximity to signal-containing nascent chains [37–41]. The eventual purification and identification of these proteins in the mammalian cotranslational system led to the discovery of TRAM (for *tr*anslocating chain-*a*ssociated *m*embrane protein [42]), the heterotetrameric *tr*anslocon-*a*ssociated *p*rotein (TRAP) complex (composed of α-, β-, γ-, and δ-subunits [43]), and the heterotrimeric Sec61 complex (composed of α-, β-, and γ-subunits [44]). In additional to being near (i.e., within cross-linking distance) to a signal sequence, each of these translocon components were also found to be functionally involved in signal sequence-mediated translocation [11, 42, 44–46].

To delineate the signal sequence interaction with the translocon with higher resolution, a site-specific cross-linking approach was combined with biochemical and genetic manipulation of the translocon components. These studies revealed several important observations. In the mammalian cotranslational system, nearly every residue in the hydrophobic core of a signal sequence was found to be adjacent to Sec61α and phospholipids [24, 47–49]. By contrast, TRAM was found to make its predominant contacts to the regions flanking the h-domain [24, 48]. For the signal sequence of prolactin, the n-domain was close to TRAM, while the α-factor signal sequence (when presented as an RNC) showed the strongest cross-links to TRAM to regions downstream of the h-domain. Although less precisely mapped, TRAPα

seems to cross-link with only longer nascent chains, potentially through residues in the mature domain [42].

In analogous cross-linking experiments using the yeast posttranslational system, translocating nascent chains were found in the vicinity of Sec62p and Kar2p (the yeast homologue of BiP) in addition to Sec61p [50, 51]. These early studies together suggested that signal sequence recognition involved a complex protein-based interaction in both the mammalian and yeast systems, but additional resolution was difficult to discern. To address this problem, the genetic manipulability of yeast was combined with site-specific cross-linking to identify the specific regions of Sec61p close to the signal sequence of a stalled translocation intermediate [18]. This systematic study found that the signal sequence of α-factor occupies a binding site surrounded predominantly by transmembrane helices 2 and 7 of Sec61p (a multispanning membrane protein containing 10 TMDs), with additional contacts to Sec62p, Sec71p, and lipid.

Because an approximately similar pattern of cross-links was observed with mammalian Sec61α for RNCs of α-factor [18], it is presumed that the principal signal-binding site formed by helices 2 and 7 is conserved across all species and modes of translocation. This is logical since homologues of the Sec61 complex form the central translocation channel in all ER and bacterial translocons, while the other accessory components (such as TRAM, the TRAP complex, Sec62, and others) are neither universally conserved nor present in every translocon. Using the cotranslational system, it has further been shown that the Sec61 complex is both necessary and sufficient for recognition of at least some model signal sequences [11]. Although definitive evidence is still lacking, cross-linking analyses suggest that TMDs interact (via their hydrophobic domain) with the same or similar site as signal sequences [52, 53]. Indeed, as for signal sequences, the Sec61 complex alone can suffice for recognition and proper orientation of at least some model TMDs of either topology [44]. Thus, signal recognition involves an essential and major interaction with a specific site within the Sec61 channel. Additional (presumably weaker, but stabilizing) interactions are specific to the translocation system: TRAM and potentially the TRAP complex are involved in the mammalian cotranslational system, while Sec62p and Sec71p participate in the yeast posttranslational system.

## VI. A Combined Framework for Signal and TMD Recognition

Considered together with the available information on protein structure, topology, and functional data, we propose the following working model for co- and posttranslational signal recognition (Figure 9.3C). In both

modes, the hydrophobic core of the signal sequence would interact with roughly the same area of Sec61α between helices 2 and 7 [18]. Indeed, even in the bacterial translocation system, these same two helices of SecY (the homologue of Sec61α) were also observed to interact with a synthetic signal peptide in detergent solution [54]. This binding site is presumably flexible (i.e., helices 2 and 7 can be moved to varying degrees) to accommodate the tremendous sequence diversity of signals (and TMDs). Furthermore, the space near helices 2 and 7 should be capable of providing a route for the signal sequence (or TMD) to access the hydrophobic core of the lipid bilayer [18, 24, 49, 52, 53, 55, 56]. Each of these conclusions derived from the mammalian, yeast, and bacterial systems is consistent with the crystal structure of an archaeal SecY complex [57]. This structure revealed that helices 2 and 7 are indeed adjacent to each other and provide a lateral exit site from the proposed pore within SecY to the lipid bilayer. Thus, it seems reasonable to conclude that in all systems, the helix 2/7 interface represents a generic hydrophobic domain recognition site that is accessed by the core region of all signals and TMDs regardless of the mode of translocation.

In the mammalian cotranslational system (Figure 9.3C, left diagram), this interaction would be stabilized by additional contacts between the n-domain and TRAM on the cytosolic face of the translocon [48]. Such contacts, although made with most or all signal sequences, would not be absolutely essential for a stable signal-Sec61 interaction for all signal sequences [11, 42, 44]. This would explain why TRAM is stimulatory for the translocation of some but not other proteins in a signal sequence-dependent manner [45], even though it interacts with substrates that do not necessarily require it for translocation. Similarly, the TRAP complex is envisioned to act similarly, but on the lumenal face of the translocon [46]. This model is consistent with the interaction between only longer nascent chains and the TRAP complex [58] and the observation that TRAP contains a large lumenal domain that appears in cryo-electron microscopy (EM) reconstructions to protrude over the lumenal aperture of the translocon [59].

Such a stabilizing role for both TRAM and TRAP would explain why both factors seem to be more important for substrates whose signal sequences are potentially weaker as judged by a noncanonical n-domain and/or a shorter and less hydrophobic h-domain [45, 46]. Furthermore, TRAM and TRAP may allow the core Sec61 complex to recognize a much broader range of signal sequences than would otherwise be possible due to constraints on the flexibility of the helix 2/7-binding site. This breadth of recognition may be particularly important in systems (such as mammalian cells) where essentially all substrates are translocated in the same mode (obligately cotranslational) instead of being subdivided between co- and posttranslational systems that involve different translocation machineries (as in yeast [6]).

A similar model of primary and stabilizing interactions may apply for posttranslational translocation in yeast (Figure 9.3C, middle diagram). Here, the translocon is composed of Sec61p, Sec62p, Sec63p, Sec71p, and Sec72p [60]. It is plausible to view the role of Sec62p and Sec71p as being designed, at least partially, for the recognition of modestly hydrophobic signal sequences that are selectively targeted to this translocon [6]. In this view, a translocon lacking these components (composed only of the Sec61 complex) could only handle signals and TMDs whose interaction with the helix 2/7-binding site is sufficiently stable to obviate additional stabilizing factors. Hence, the cotranslational translocon in yeast (which lacks homologues of TRAM and the TRAP complex) would accommodate only the limited subset of signals that could be accommodated into its principal helix 2/7-binding site (Figure 9.3C, right diagrams). This constraint may explain why in yeast, there exist two different cotranslational translocons composed of either Sec61p or its homologue Ssh1p [61]: each may be capable of recognizing different subsets of signal sequences and operate in parallel to together accommodate a wider range of substrates [62].

Thus, in yeast, the problem of accommodating the remarkable diversity of signal sequences seems to have been solved by the use of multiple parallel translocation machineries that each has distinct (but perhaps overlapping) substrate-specificity. By contrast, higher eukaryotes (such as the mammalian system) may have overcome this same problem by evolving translocon-associated factors such as TRAM and TRAP to stabilize weakly interacting signal sequences that otherwise could not be efficiently recognized by the Sec61 complex alone. Such a qualitatively different solution (of multiple weak interactions colluding to accomplish the final outcome) to the signal recognition problem may have significant implications for how translocation could be regulated by the cell (see the final section of this chapter).

The same basic mode of signal-translocon interaction is likely to be applicable for the initial recognition of TMDs during translocon engagement (Figure 9.3B; compare to Figure 9.3A). Because the hydrophobic domain is usually both longer and more hydrophobic than signal sequences, stabilizing interactions with TRAM and TRAP may be less critical for initial TMD recognition and engagement of the translocon (although this remains to be tested experimentally). Even if initial recognition is less dependent on accessory factors, the problem of correct orientation could very well involve interactions with accessory components of the translocon. This seems entirely plausible in the case of TRAM and TRAP because these components seem to interact with regions flanking the hydrophobic domain of a signal sequence [24, 42, 47, 48] and may similarly associate with flanking domains of a TMD. As with signal sequences, features intrinsic

to Sec61 may simply be insufficient to fully accommodate the diversity of TMDs (especially in the mammalian system) with which it must interact. Thus, accessory components (while dispensable for the few model TMDs thus far studied [44]) may be particularly important for assimilating the various parameters of the substrate (including hydrophobicity, charge of flanking domains, length, and folding of flanking domains) into a final unique topologic orientation.

The sequence features of a signal or TMD that determine the need for components in addition to the Sec61 complex are not well studied in either the mammalian or yeast systems. Furthermore, the mechanisms by which such accessory components like TRAM, TRAP complex, or Sec62p facilitate signal recognition remain unknown. Finally, whether yet other components are also involved in substrate-specific aspects of signal recognition is also not known. Numerous proteins, particularly in the mammalian system, have been identified to be at or near the site of translocation (reviewed in [3, 63]). These include proteins with known functions (such as the multiprotein oligosaccaryl transferase complex or five protein signal peptidase complex), as well as many others whose functions are not known (including for example p180, Mtj1, RAMP4, and p34). While none of these are absolutely essential for translocation of at least the simplest model substrates [44], it is not clear whether they play essential or stimulatory roles in translocation of select substrates. As was exemplified by the TRAP complex [46], the functional role(s) of such accessory factors in translocation may elude detection until the proper substrate(s) is examined.

## VII. Gating of the Protein-Conducting Channel of the Translocon

In addition to discrimination from nonsecretory proteins, signal sequence recognition is an important event in preparing the translocon for subsequent protein translocation. Hence, the signal recognition step is thought to be coincident with a change in the translocon that leads to its opening toward the ER lumen [11, 13], a step termed translocon gating. Although it is clear that this is an essential step in initiating translocation, the mechanism by which it occurs remains unknown. Presumably, the initial binding of a signal sequence (or TMD) to the Sec61 complex (likely at the helix 2/7 interface) triggers a conformation change in either the Sec61 complex itself and/or associated components to convert the translocon from a ''closed'' to ''open'' configuration. Understanding the mechanism by which this occurs is intricately tied to the general issue of membrane permeability: the translocon should normally remain closed when it is inactive yet open

during substrate translocation such that in neither situation can small molecules easily pass across the membrane.

At present, the question of how the membrane permeability barrier is maintained during different stages of protein translocation remains a matter of considerable debate. It is clear, however, that resolving this issue will require information about both the architecture of the translocon, the structure of its individual constituents, and how they are assembled and changed during the functional translocation cycle. This will provide critical information about the nature of the translocation pore, its size, how it might be opened and closed, and how its permeability to small molecules can be controlled both during and in the absence of substrate translocation. At present, such structural and organizational information about the translocon and the pore are only beginning to emerge, leaving the mechanism of membrane permeability maintenance unresolved.

The first experimental studies to begin addressing the issues of pore size and membrane permeability were in the mammalian cotranslational system. Translocation intermediates containing a fluorescently labeled amino acid were used as a probe of both the environment surrounding the nascent chain [36] and the accessibility of this environment to exogenously added fluorescence quenchers [13, 64–66]. The ability to control substrate length (and hence, the stage of translocation), the position of the probe, and the size and location of the fluorescence quenchers allowed various parameters of the translocon to be deduced. From these studies [3], the pore sizes of inactive versus engaged translocons were measured to be ~8–10 and ~40–60 Å, respectively [64]. Preventing the passage of small molecules through this pore depended on alternately sealing the channel with either a ribosome on the cytosolic side or BiP on the lumenal side [64–66]. Sequences in the nascent polypeptide are proposed to choreograph the dynamics of channel gating by the ribosome and BiP to allow substrate transport without small molecule leakage [66]. An electrophysiological approach also suggested that purified Sec61 complex in lipid bilayers may contain pores as large as 60 Å that can be blocked by BiP [67].

Although the model derived from the fluorescent probe approach is internally consistent and compatible with many other biochemical experiments in the mammalian cotranslational system, several arguments against it have been raised. In one experiment, the inability to detect folding of even a small domain while it is inside the translocon [68] seemed at odds with the proposed 40- to 60-Å pore size [64, 67]. However, it is not clear how generalizable the results from either approach are since in each case, a single (and different) substrate has been examined to measure pore size. In other experiments, structural studies using cryo-EM of RNCs bound to the translocon failed to see a tight seal between the ribosome

and translocation channel that was expected from the fluorescence quenching studies [69–72]. However, an *inability* to see density by cryo-EM can be difficult to interpret since it could be due to increased flexibility in those regions of the structure, loss of ancillary translocon components on solubilization and sample preparation, or sample heterogeneity. Thus, cytosolic or membrane components in addition to the ones visualized by cryo-EM may form the putative seal between the ribosome and membrane. Indeed, several abundant membrane components have been identified associated with the translocon (some with large cytosolic domains such as p180) [73] whose functions remain unclear. Thus, there are some potentially plausible ways to reconcile much of the seemingly conflicting data gathered on membrane permeability and translocon architecture of the mammalian cotranslational system.

More problematic, however, is the argument that the proposed mechanism involving the ribosome and BiP during mammalian cotranslational translocation does not shed light on how the permeability problem is solved in other modes of translocation or in bacterial systems. In the posttranslational pathway, the ribosome is not involved in translocation, precluding a role for it in maintaining the permeability barrier. In bacteria, it is unclear what would serve the function of the lumenal gate proposed for BiP in the mammalian system. Because of these difficulties, a more generally applicable and evolutionarily conserved solution to the permeability barrier problem has been sought. The most insight into such a putatively conserved mechanism comes from interpretation of the recent high-resolution crystal structure of an archaeal SecY complex [57, 74].

In this structure, a single SecY complex was found to form a channel-like structure with a very small pore flanked on the lumenal and cytosolic sides by funnels. The narrow constriction between these two funnels is only ~5–8 Å in diameter and lined by several hydrophobic residues that together form the ''pore ring.'' If the channel formed by a *single* SecY complex is the functional pore through which the substrate is transported, the small size and flexibility of the ''pore ring'' side chains would then form a relatively snug fit around a translocating polypeptide. This mechanism of translocation would solve the permeability problem because the nascent chain itself can occlude the channel during translocation. Furthermore, another small segment of the SecY protein (termed the ''plug'' domain) appears to occlude the pore in its inactive state [57]. Thus, no additional components would be required to maintain permeability except the Sec61/SecY complex, which forms the channel in all modes of translocation.

Mutagenesis studies have demonstrated movement of the plug domain on initiation of translocation, supporting its proposed role in gating [75]. However, other studies question an essential role for the plug domain [76]

since even its complete deletion was nonetheless compatible with viability in yeast (although various translocation defects were noted). In addition, cross-linking between nascent chains and residues near the pore ring (but not elsewhere) supports the polypeptide transit path proposed on the basis of the SecY structure [77]. Based on these and other considerations, it has been argued that a hydrophilic pore is not formed at the interface of multiple SecY complexes, but within a single SecY whose permeability is maintained by its plug domain (reviewed by [4, 74]. Whether this model proves to be true in all modes and systems of translocation, and hence universally explains the permeability problem, remains to be investigated by high-resolution structures of translocons engaged in active translocation. At present, computational fitting of the archaeal X-ray structure into a cryo-EM structure of an RNC-engaged *Escherichia coli* translocon suggests that two SecY molecules may ''face'' each other with communicating pores [78]. If validated, this may suggest that the basic unit of translocation can be reorganized significantly by associating components (like the ribosome) when changing between inactive and active states.

Hence, in eukaryotic systems, the basic unit of translocation may have evolved into a more malleable oligomeric structure in which the pores of multiple Sec61 complexes can indeed be combined to form a larger translocon that changes to meet the demands of the substrate. This explanation would necessitate additional protein complexes that facilitate this reorganization and new mechanisms to solve the permeability problem. This view could help to reconcile the fluorescence data in the mammalian system with alternate models of gating derived from other systems. While this might seem unnecessarily complicated, it is not unreasonable given the existence of numerous eukaryotic-specific translocation components whose functions remain largely unknown (such as Sec62, Sec63, TRAM, or TRAP, among many others).

At present, the choice among the different views depends largely on where a philosophical line is drawn. On the one hand is the tremendous degree of evolutionary conservation of the most fundamental features of protein translocation that has allowed information across multiple kingdoms to be combined into explanations applicable to all systems. On the other hand is the equally powerful feature of evolution to forge new biological principles using the same basic constituents. Clearly, the former is justified when one considers examples such as the SRP pathway, while the latter is strikingly exemplified by the evolution in eukaryotes of mechanisms to ''pull'' nascent chains across the membrane from a system initially designed to ''push'' such chains from the cytosolic side. Ultimately, experimental results will be needed to resolve these issues and determine the degree to which evolution has been conservative versus inventive in shaping eukaryotic protein translocation across the ER.

## VIII. The Energetics of Protein Translocation

Regardless of the structural features of the translocon pore or the precise mechanism of translocon gating, it is clear that after this event has occurred, the nascent polypeptide spans the membrane bilayer through the aqueous channel formed by the Sec61 complex [13, 47]. Hence, a portion of the nascent chain can access the ER lumenal environment, while the remainder of the chain has yet to be translocated. The channel within which the nascent chain resides does not appear to interact with the chain; it is instead thought to be a relatively inert passive conduit. Indeed, early studies artificially releasing a nascent chain stalled in its cotranslational translocation showed it to be capable of bidirectional movement to either the lumenal or cytosolic side of the membrane [79]. This suggested that unidirectional vectoral movement of the nascent chain must be imparted by accessory factors and presumably requires the input of energy.

In the cotranslational pathway, the substrate is thought to be ''pushed'' across the membrane by the ribosome (Figure 9.4A). In this model, the

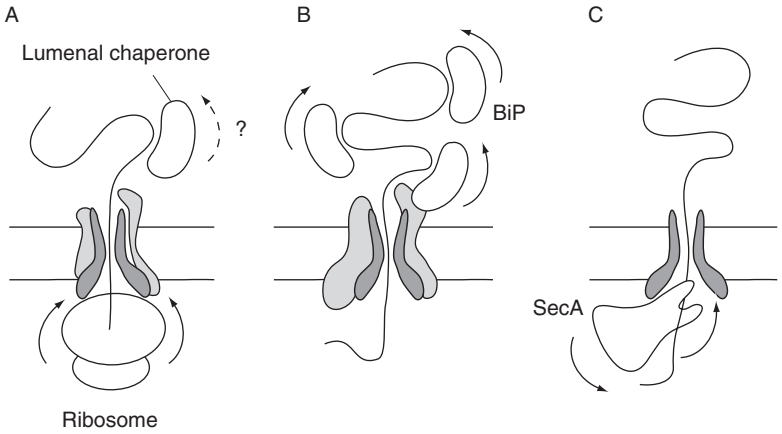



FIG. 9.4. Energetics of protein translocation. Energy from several sources is harnessed for vectorial transport of substrates in the different translocation systems. In cotranslational translocation (panel A), the architecture of the ribosome-translocon complex allows translocation to occur concurrently with polypeptide elongation, thereby harnessing the energy of protein synthesis for transport. In posttranslational translocation (panel B), the lumenal chaperone BiP utilizes ATP hydrolysis for repeated rounds of binding and release from translocating polypeptides to effect transport by a ''molecular ratchet'' mechanism. The recruitment of BiP to the translocon and regulation of its ATPase cycle are provided by Sec63. BiP may also play a role in cotranslational translocation, although this remains to be fully investigated. In bacterial systems (panel C), a cytosolic ATPase termed SecA seems to act as a motor that ''pushes'' nascent chains through the translocon. A comparable system in eukaryotes has not been described.

pushing is critically dependent on the architecture of a ribosome-translocon complex that provides a continuous path from the peptidyl transferase site within the ribosome to the lumenal aperture of the translocon. When configured this way, the nascent chain is thought to essentially have only one path of transit as it is elongated by continued translation. Thus, the energy of protein synthesis is simultaneously used to support translocation. This view is supported by the demonstration of a contiguous path through the ribosome-translocon complex [13], the apparent alignment of conduits through both structures [69], the shielding of translocating nascent chains from the cytosol [11, 13], and the reconstitution of successful translocation with the Sec61 complex as the only translocon component in the membrane [11, 44]. Thus, for at least some substrates, cotranslational translocation can proceed solely on the basis of energy expended for protein synthesis.

However, several observations suggest that this model of cotranslational translocation is insufficient to explain the vectoral transport of all cotranslationally translocated substrates. First, the putative tight seal between the ribosome and translocon that precludes nascent chain slipping during translocation may not be uniformly maintained for all substrates. In some instances, nascent chains have been shown to be accessible to the cytosol during translocation [80, 81]. In other cases, the nascent polypeptide did not become shielded from the cytosol shortly after its docking at the translocon, instead being cytosolically accessible for prolonged periods early in translocation [82]. Cytosolic accessibility is also predicted on the basis of a relatively large gap observed between the ribosome and translocon in cryo-EM structures [59 69–72, 83], and the fact that cytosolic loops in multispanning membrane proteins must have some means of escaping the ribosome-translocon tunnel during translocation. Thus, it is reasonable to conclude that many nascent chains have access to the cytosol at one or another point during their translocation, and therefore at least have the potential to slip into the cytosol rather than being translocated.

This potential to slip may explain why some studies have found that nascent chain-binding proteins on the lumenal side of the membrane can stimulate cotranslational translocation [84–88]. It is plausible that such additional factors are required to bias transport only in some circumstances or for some substrates, perhaps explaining why this lumenal factor requirement has not been uniformly observed by all investigators. Additional studies systematically and quantitatively examining the requirements for cotranslational translocation of many different types of substrates will be needed to fully identify all of the factors that impart unidirectionality to the transport process.

In the yeast posttranslational pathway (Figure 9.4B), the ribosome and energy of protein synthesis cannot be exploited for translocation. Thus,

other factors, such as the lumenal chaperone Kar2 (the yeast homologue of BiP), are obligately required for vectoral translocation [60, 88–90]. Kar2 is recruited to the lumenal side of the translocation site via its interaction with Sec63, a J-domain protein that regulates the ATPase activity of BiP [89, 91]. At the site of translocation, Kar2 binds to the substrate and prevents its back sliding to the cytosol [60, 89, 90]. Repeated rounds of ATP-dependent binding and release, coupled with Brownian motion of the nascent chain, are though to drive substrate translocation into the ER lumen by a ''molecular ratchet'' mechanism [92]. This role of BiP/Kar2 in the translocation of proteins across the ER membrane is discussed in more detail in Chapter 10, this volume.

A yet different translocation mechanism is involved in posttranslational translocation in bacteria (Figure 9.4C). SecA, a dissociable subunit of the translocon, appears to harness its ATPase activity to ''push'' substrates in a stepwise manner through the SecY pore. In addition, the proton motive force can drive translocation after the initiation by SecA and can be continued and completed even after depletion of ATP [4, 93]. This translocation system (described in detail in Chapter 2) further underscores the basic idea that components of the membrane-embedded translocon or translocation channel do not directly provide any intrinsic directionality to transport; instead, the energetics and mechanisms for directionality are imparted by reversibly associated accessory factors. The reason for this ''division of labor'' is not clear, but may allow translocons to be adapted for multiple purposes (such as forward translocation and retrotranslocation) simply by changing the spectrum of associated components.

## IX. The Biogenesis of Membrane Proteins

For secretory and many single-spanning proteins, the issues thus far considered (recognition, engagement of the translocon, and vectoral transport) largely suffice to explain their translocation. However, multispanning membrane proteins have several TMDs that must each be recognized, oriented appropriately, laterally inserted into the membrane bilayer, and assembled with other TMDs to create the final product. All of these events are thought to occur cotranslationally in all systems ranging from bacteria to mammals. The reason may have to do with the relatively intractable problem of even transiently maintain a highly hydrophobic protein containing several TMDs in a configuration capable of subsequent translocation. Instead, cells appear to deal mostly with TMDs sequentially, as they emerge from a ribosome docked at an already engaged translocon. The mechanism

involved in the handling of internal TMDs is very poorly studied, and remains largely in the realm of speculative working models.

It is reasonable to assume (albeit with little experimental data) that recognition and orientation of internal TMDs operates by principles similar in many ways to the initial recognition event of signal sequences and signal anchor sequences (Figure 9.5). The one obvious exception is that the orientation of preceding TMDs already embedded in the membrane is likely to impose rather strict constraints on the topology that can be achieved by subsequent TMDs. Unless previously synthesized TMDs are

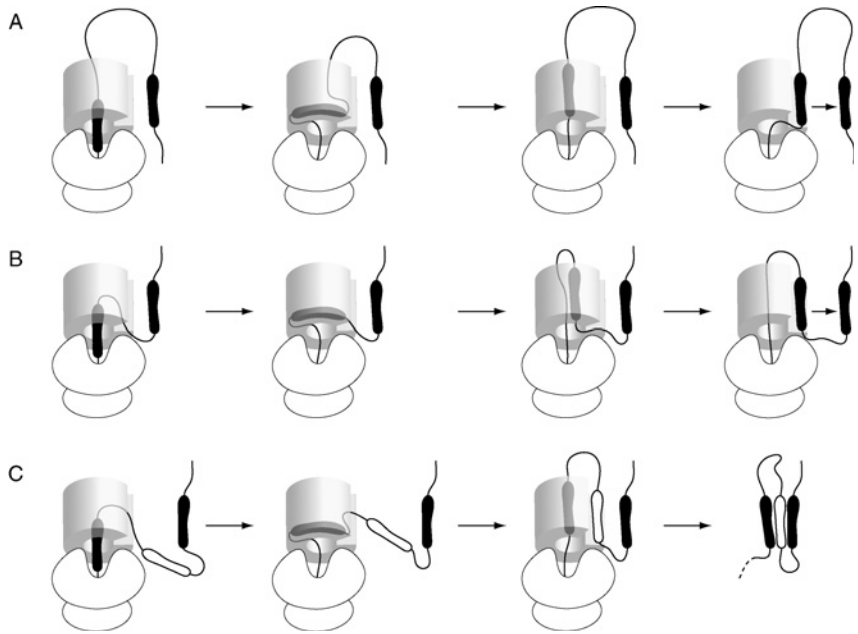



FIG. 9.5. Recognition and membrane integration of internal TMDs. When a TMD within the ribosomal tunnel approaches the translocon (left diagram in each panel), it appears to acquire secondary structure [94] and may induce conformational changes in the translation and translocation machinery such that the channel is closed to further translocation (not depicted). The TMD is subsequently recognized by the translocon (second diagram in each panel) in a step that is poorly understood by may be similar to the initial signal/TMD recognition step (see Figure 9.3). Depending on constraints such as the topology of preceding TMDs, the charge distribution flanking the TMD, length, and hydrophobicity, the TMD is oriented within the translocon (third diagram in each panel). In some cases, these constraints may force a moderately hydrophobic or nonhydrophobic segment of the polypeptide (white segment in panel C) to acquire a membrane-spanning configuration. In the final step, the TMD is laterally moved into the lipid bilayer, either on its own or perhaps in conjunction with preceding TMDs with which it assembles.

reoriented after their initial handling (an event for which some precedent exists), the incoming TMD must acquire a topology opposite to the preceding TMD (Figure 9.5A and B). In the context of this constraint, how then are cotranslationally synthesized internal TMDs recognized, oriented, integrated, and assembled?

Quite surprisingly, fluorescence probes incorporated into TMDs suggest that they may first be recognized while still inside the ribosome [66]. Not only does the TMD begin to acquire secondary structure inside the ribosomal tunnel, but this recognition seems to induce changes in the translocon machinery to which the ribosome is bound [94, 95]. In particular, the translocon becomes sealed (i.e., a closed configuration) on the lumenal side, while the ribosome-translocon interaction changes in a way that allows nascent chain exposure to the cytosol [96, 97]. This translocon configuration (sealed from the lumen and open to the cytosol) is similar in many ways to the configuration encountered by a signal sequence or TMD when it is first targeted to a vacant translocon. It is therefore appealing to consider a model in which every TMD, regardless of its position in a multispanning membrane protein, encounters the translocation apparatus in essentially the same way. In this view, the translocon is effectively ''reset'' to a baseline configuration before the emergence of every TMD.

From this baseline configuration, we would imagine that the translocon recognizes and orients each successive TMD by using similar parameters as those used for signals and signal anchors. This conclusion is consistent with the observation that the charge distribution flanking internal TMDs is similar (although not quite as obvious) to that of first TMDs [28]. These same recognition parameters, together with the additional constraint of previous TMDs, stabilize the intended orientation of most TMDs. Because many TMDs may not be as robustly oriented on their own as others, the constraint of previous TMDs may be more important in some instances than others. Indeed, experiments intentionally altering the flanking charges of internal TMDs have shown that TMDs can be ''forced'' into specific orientations solely on the basis of surrounding TMDs (Figure 9.5C). For example, a TMD that is essentially incapable of engaging the translocon on its own could potentially be forced to span the membrane by two very strongly oriented flanking TMDs [98, 99]. Such unusual insertion mechanisms may be especially important for the biosynthesis of many membrane proteins (such as ion channels) whose internal TMDs may need to be relatively nonhydrophobic for correct function of the final product.

Considering the tremendous range of TMDs that are handled by the translocon, it seems unlikely that a single mechanism is used for their uniform recognition and orientation. Instead, a combination of selfcontained sequence elements (e.g., hydrophobicity, flanking charges), constraints

imposed by other TMDs, cooperation between adjacent TMDs [56, 100], and even reorientation mechanisms [101] all contribute to complex membrane protein biogenesis. This multifactorial aspect of TMD integration may help to understand why prediction of topology is nontrivial, and how many regions of limited hydrophobicity can nonetheless serve as TMDs.

## X. Lateral Exit of TMDs from the Translocon

After recognition and orientation of a TMD by the translocon, it must be moved from the aqueous translocation channel to the hydrophobic lipid bilayer. Several nonmutually exclusive mechanisms have been proposed to explain how this lateral exit occurs (reviewed in [3, 102, 103]). It is likely that each of these mechanisms operates under different circumstances, depending on the features of the TMD and other parts of the substrate (Figure 9.6).

Prior to integration, cross-linking studies have shown that the TMD is adjacent to Sec61α, lipids, and, in some cases, TRAM [52, 53, 55]. Because cross-linking probes positioned in adjacent sites within a TMD show different efficiencies of cross-linking to Sec61α and TRAM, it is thought that the TMD is bound to a specific site in the translocon [95]. This site may be the same as the signal sequence-binding site involving helix 2/7 of Sec61α. Binding to the translocon is also presumably important to prevent further translocation into the lumen during continued translation. From this site, the TMD then moves into the lipid.

The simplest mechanism of lateral TMD movement involves its partitioning via its hydrophobic core into the lipid bilayer [52]. Because the TMD, even when in the translocon, can be cross-linked to lipids, it is presumably already positioned at the lateral exit site. The model is that on further translation, the tether keeping the TMD in proximity to the translocon is lengthened, allowing the TMD free reign to access its energetically favored environment. Hence, the TMD moves from the translocon to the lipid bilayer before protein synthesis had terminated (Figure 9.6A). Consistent with this partitioning model, introduction of charges into the TMD slowed its movement into the lipid bilayer [52]. Furthermore, this mechanism (at a minimum) appears to involve only the Sec61 complex since integration by partitioning has been reconstituted using purified components [52].

In another set of studies with different substrates, the TMD was observed to remain within cross-linking distance to Sec61α and/or TRAM for a prolonged time, up until the terminal codon was reached by the ribosome [55, 95]. This has been interpreted as a mode of integration

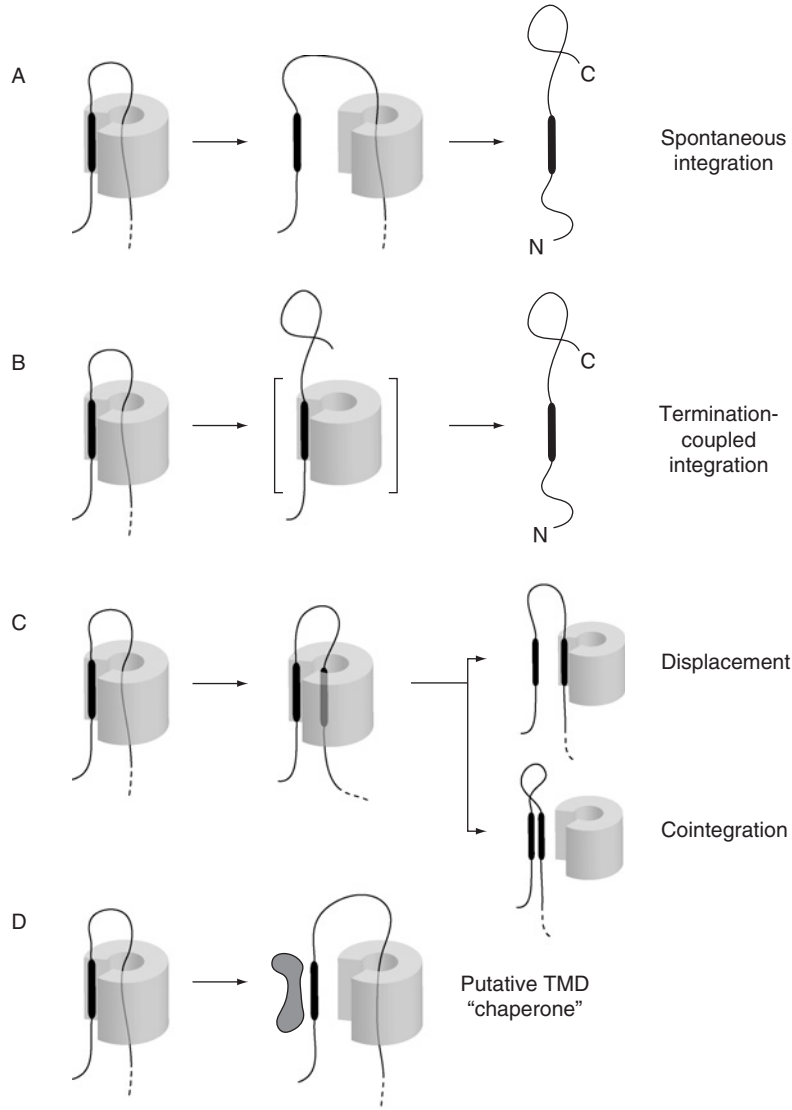



FIG. 9.6. Lateral exit of TMDs from the translocon into the lipid bilayer. After recognition and orientation of a TMD in the translocon (see Figure 9.2 and Figure 9.5), it is thought to reside at a lateral exit site at the interface between the channel and lipid bilayer (possibly between helix 2 and helix 7 of Sec61α; see Figure 9.3C). Each of the several mechanisms of TMD movement into the lipid bilayer that have been suggested is indicated in panels A through D. See text for additional details.

being distinct from the partitioning model. It is possible that the conformational change in the ribosome-translocon complex on termination is the critical event in driving TMD integration (Figure 9.6B). The mechanistic difference between integration during synthesis versus integration on termination is not clear, but could have to do with different substrates or experimental conditions. Given that most detailed studies of TMD integration are performed using artificially designed substrates, it is not yet obvious how commonly each mechanism is used in naturally occurring membrane proteins.

The third mechanism of integration involves ''displacement'' of a TMD from its binding site in the translocon by a subsequent TMD (Figure 9.6C). In one version of this view, only one principal binding site for a TMD (or signal sequence) exists within the translocon. Hence, each subsequent TMD would displace the previous TMD from this site, forcing its integration into the lipid bilayer [104–106]. A variation of this view is that some TMDs are not sufficiently stable on their own in the lipid bialyer, and are thus held at a protein–lipid interface until it can cointegrate with another TMD (with which it interacts). Indeed, examples of membrane protein integration where two TMDs cooperate to facilitate each other's integration have been described [56, 100]. Extrapolation of this line of thought leads to a model where the translocon (perhaps in combination with associated factors) can retain multiple TMDs (or retrieve them to the translocon ever after integration) to facilitate the assembly of three or more TMDs cotranslationally before their en bloc release into the lipid bilayer [105, 107].

And finally, although poorly studied, complex membrane protein assembly events such as multiple TMD assembly may require chaperones specialized for handling TMDs (Figure 9.6D). TRAM, an importin-like ER protein termed importin$\alpha$16, and a yet unidentified factor termed PAT10 have each been suggested to serve in this role on the basis of cross-linking studies [52, 55, 104, 108]. At present, experiments examining the consequences of removing any of these components for membrane protein integration remain to be performed. Thus, their functional roles in the integration process remain largely a matter of speculation. In fact, aside from the Sec61 complex, no other component has been functionally shown to directly influence integration (defined as the lateral movement of a TMD from the aqueous translocon to the lipid bilayer). Setting up robust and well-defined assays for TMD orientation, integration, and assembly in a system readily amenable to selective depletion of individual translocon components will be an important goal in defining the mechanisms of membrane protein biogenesis.

## XI. Regulation of Protein Translocation

The evolution of a complex endomembrane system in eukaryotes provides several advantages to the cell, some of which are more obvious than others. These advantages include increased capacity, quality control, quantity control, and regulation. On translocation into the ER, a protein is still available to a eukaryotic cell before its secretion or exposure at the cell surface; by contrast, translocation is largely synonymous with exit from the cell in prokaryotes. This intracellular availability prior to secretion or presentation on the cell surface has been thoroughly exploited to confer several important advantages to eukaryotes.

The most important advantage is the opportunity for quality and quantity control: since a translocated protein in eukaryotes is not lost to the extracellular space, there is time to impose a ''recall'' in instances where the protein is not desired. Hence, if a protein is not matured or assembled properly, it is rerouted for degradation (i.e., quality control) [109, 110], thereby avoiding the potentially detrimental consequences of misfolded or incomplete secretory and membrane proteins. This has almost certainly facilitated the evolution of very complex secretory proteins (such as apolipoprotein B) or multicomponent membrane protein complexes (such as the T-cell receptor). Similarly, regulated degradation during or shortly after translocation allows the abundance of secretory or membrane proteins to be modulated in response to need (i.e., quantity control, exemplified by HMG-CoA reductase [111] or apolipoprotein B [112, 113]).

Furthermore, the intracellular compartmentalization of secretion allows secretory and membrane proteins to be stored until they are needed [114], at which point they can be rapidly delivered to selected regions of the cell surface by exocytosis. Thus, secretion of extracellular proteins or surface expression of membrane proteins can be rapid, quantal, and temporally and spatially regulated. These examples illustrates an important general principle: the disadvantages of increased cost and lower efficiency of a more complex, multistep process (e.g., the secretory pathway) can be offset by the benefits of a greater degree of regulatory control. Thus, potentially regulatory aspects of the secretory pathway are likely to be most thoroughly developed in systems where control, and not just energetic cost, is of the utmost importance.

In which organisms is the highest premium placed on precise control of secretory and membrane protein biogenesis? The answer is multicellular organisms, whose fitness depends not only on the health of individual cells but equally (or perhaps even more) on the ways those cells interact, communicate, and function as complex units. Such communication and

interactions are intimately dependent on secreted and cell surface proteins whose amounts at the right time and place must be carefully regulated. Thus, completely healthy *individual* cells in a complex organism can nonetheless lead to failure of the organism if they do not function coordinately in extremely precise ways. Countless examples of this idea can be found in human physiology and disease, including the regulation of blood pressure, reproductive cycles, stress, appetite, and weight regulation. It should, therefore, come as no surprise that *each and every step* in the secretory pathway that has been examined was discovered to be regulated to tightly control the levels of secretory and membrane proteins in response to cellular and organismal needs. Will protein translocation prove to be any different once more complex (and subtle) aspects of this process have received experimental attention? Almost certainly not.

How then one might conceptualize a framework for translocational regulation that can guide future investigation? At the outset, it is instructive to consider analogies to other regulatory systems for common themes that can be applied to translocation. In this vein, a grossly simplified discussion of transcriptional promoters and their regulation is useful [115–118] (although similar arguments can be made equally well with any other regulatory process). In transcription, sequence features that are common to all promoters are accompanied by sequence elements that are unique to each individual promoter [119]. Thus, each promoter is unique, but contains at least some common elements that allow it to be recognized as a promoter per se. The common elements allow a core (or ''general'') machinery to mediate transcription [116, 117], while the unique elements impose requirements for additional machinery that regulate the recruitment or activity of the core components [115, 118]. The combinatorial expression or modification of the unique machinery can dramatically influence the activity of any given promoter. By regulating individual components of the unique machinery in a temporal or cell type-specific manner, transcriptional regulation of individual promoters can be achieved independently of each other. Thus, sequence diversity of promoters combined with diversity in the components that recognize them allows selective regulation of genes that all nonetheless use a commonly shared core machinery for transcription.

Applying this general idea to translocation allows at least one mechanism of regulation to be conceptualized. Here, signal sequences are viewed as loosely analogous to promoters, and the evolutionarily conserved components of the translocation machinery (i.e., SRP, SR, and Sec61 complexes) are analogous to the core transcriptional machinery. Signal sequences are indeed extremely diverse, with each substrate containing an effectively unique signal, while nonetheless sharing certain common, recognizable features [25, 120]. The common features of the signal appear to be the

elements that are recognized by the core machinery, such as SRP54 and the Sec61 complex. The unique features of the signal appear to impose additional constraints on signal function by requiring the presence of additional factors at the translocation site such as TRAM or the TRAP complex [42, 45, 46]. These additional components can be modified (e.g., by phosphorylation [121–123]), which potentially may selectively modulate their activity (although this has yet to be examined). Thus, even using only the limited information that is currently known, one can easily envision the basic elements of a substrate-specific system of translocational regulation (Figure 9.7): (1) diversity in structure and function of signal sequences that share a bare minimum of common features; (2) diversity in ''accessory'' components that influence recognition by a core translocation machinery of some but not other signals; and (3) selective changes in expression or modification of the ''accessory'' components that could affect the outcome of translocation for some but not other substrates.

This view of regulating translocation by the combinatorial functions of accessory components can be readily expanded to incorporate the many other factors at or near the site of translocation whose functions remain elusive. In the mammalian system, these include Sec62, Sec63, p180, p34, Mtj1, RAMP, a TRAM homologue, and yet unidentified proteins observed by cross-linking studies. Each of these components could potentially play stimulatory (or inhibitory) roles in the translocation of selected substrates, with the specificity encoded in the sequence diversity of the signal. Such accessory components can not only be modified, but themselves regulated at steps such as alternative splicing [124] or differential expression [125] to influence their function. Thus, there exist more than enough sources for modulatory activities to theoretically provide exquisite specificity in the regulation of signal sequence function, and hence translocation. Proof-of-principle that differences in signal sequences among substrates can indeed be exploited to selectively modulate translocation comes from the recent discovery of translocational inhibitors [126, 127]. These molecules appear to work by selectively inhibiting the interaction between Sec61 and some but not other signal sequences [126]. Thus, signal sequence function can be selectively, potently, and reversibly modulated in trans.

Initial evidence that protein translocation can indeed be physiologically modulated in a substrate-selective, cell-type specific way has been provided by quantitatively examining the efficiency of signal sequence function *in vivo* [128]. Not only were different signal sequences found to have different efficiencies within a given cell type, but they also varied independently in a cell type-specific manner. For example, one signal sequence was observed to be significantly more efficient than another signal in a particular type of cell; however, in a different cell type, the two signals were found

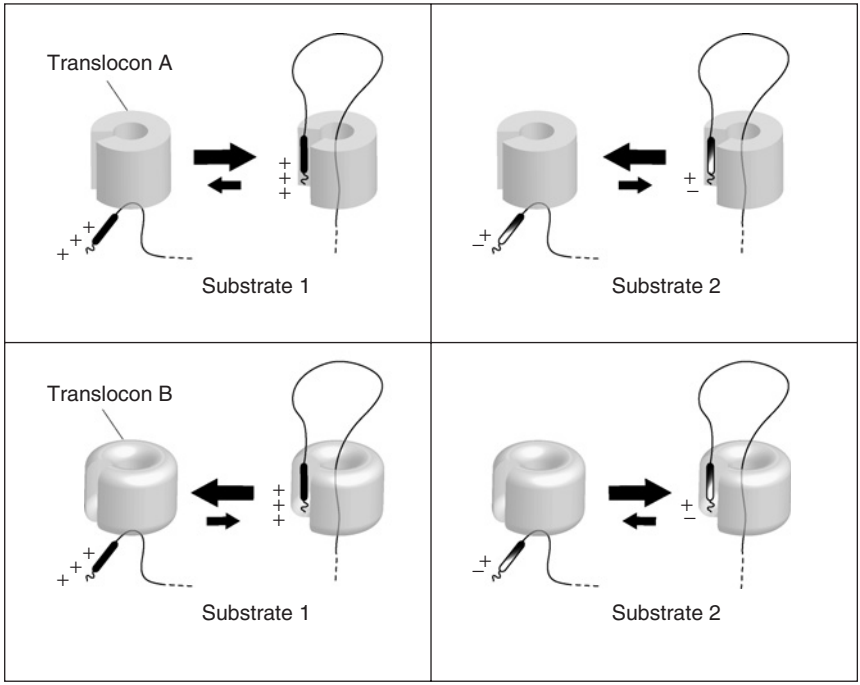



FIG. 9.7. Potential mechanisms for selective translocational regulation. Translocons in the ER of mammalian cells are capable of residing in multiple states. These different states might be distinguished by the presence or absence of various accessory factors (like TRAM, TRAP, Sec62, Sec63, Mtj1, p180, RAMP4, and others), modifications such as phosphorylation [121–123], or use of alternative splice variants or homologues of key components. Hypothetical translocons in two different states (A and B) are depicted in the top and bottom diagrams. Signal sequence recognition by these two translocons is speculated to be different with respect to the efficiencies with which they position signals of different features. For substrate 1, whose N-domain is highly basic, translocon A efficiently positions the signal in the ''looped'' orientation and hence favors translocation (top left). By contrast, the same substrate is not efficiently recognized by translocon B (bottom left). For substrate 2, whose signal sequence is different than substrate 1, efficient recognition is mediated by translocon B (bottom right), but not translocon A (top right). Thus, the differences among signal sequences combined with the malleability of translocon states would allow substrate-specific regulation of translocation into the ER.

to be equally inefficient. Similarly, the relative efficiencies of different signal sequences appeared to change even within a population of cells when assessed at different stages of growth.

In another series of experiments, it has recently been shown that translocation efficiencies of selected substrates can be responsive to environmental stimuli. For example, during acute ER stress, the translocation of

some but not other proteins was shown to be attenuated to varying degrees in a signal sequence-selective manner [129]. These aborted translocation products were routed for degradation by the cytosolic proteasome system. One purpose for this rerouting from a translocated to degradative fate (a process termed ''preemptive quality control'' or pQC) may be to protect the ER lumen from excessive protein misfolding during ER stress. Indeed, forcing the prion protein (PrP) to be constitutively translocated under these conditions by use of a highly efficient signal sequence caused increased sensitivity to ER stress and increased PrP aggregation in the ER lumen. Thus, the entry of proteins into the ER is not necessarily a constitutive process predestined by the sequences of the substrate. Rather, it is dependent on and potentially regulated by a translocation machinery that may be responsive to changes in cellular conditions.

Modulating translocation not only provides a mechanism for quantity control (i.e., the ability to change the abundance of the protein in the secretory pathway), but may also be a means to generate alternative forms of certain proteins that reside in another compartment (where it could potentially serve a second function). Examples of proteins that may have such alternative functions in different compartments have been suggested (summarized in [32, 128]). At present, at least one example of the physiological relevance of such an alternatively localized population has been provided by analyses of the ER-lumenal chaperone Calreticulin (Crt). Here, the signal sequence of Crt was shown to permit a small but detectable percent of the total to fail in its translocation and reside in the cytosol [130]. Remarkably, forcing efficient translocation by use of another signal sequence influenced gene expression mediated by the glucocorticoid receptor, a function previously ascribed to cytosolic Crt.

The degree to which translocational regulation is beneficially utilized for the generation of functional diversity or quantity control of secretory pathway proteins remains to be investigated and represents a largely unexplored area of protein translocation. It is clear that in addition to developing a working framework for the plausible ways that translocation might be regulated (e.g., as in Figure 9.7), it will be important to identify additional tractable model systems. Whereas the study of essential and constitutive facets of translocation has required simple and highly robust model systems, the study of regulation will probably necessitate more complex substrates and potentially new experimental methods. Furthermore, as in other fields, the consequences of misregulation may be more nuanced than defects in basic translocation. For example, mice disrupted for the translocon accessory component RAMP4 display an ER stress-related phenotype [131]. In addition, humans containing a mutant Sec63 develop polycystic liver disease [132, 133]. Hence, the study of translocational regulation

may require analyses in more complex organisms and systems (such as *Caenorhabditis elegans*, *Drosophila*, or mouse models) than have yet to be employed in this field. These areas of study represent challenging but physiologically important directions for the future.

## References


1. Wickner, W., and Schekman, R. (2005). Protein translocation across biological membranes. *Science* 310:1452–1456.
2. Schnell, D.J., and Hebert, D.N. (2003). Protein translocons: multifunctional mediators of protein translocation across membranes. *Cell* 112:491–505.
3. Johnson, A.E., and van Waes, M.A. (1999). The translocon: a dynamic gateway at the ER membrane. *Annu Rev Cell Dev Biol* 15:799–842.
4. Osborne, A.R., Rapoport, T.A., and van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. *Annu Rev Cell Dev Biol* 21:529–550.
5. Shan, S.O., and Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. *FEBS Lett* 579:921–926.
6. Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* 134:269–278.
7. Hatsuzawa, K., Tagaya, M., and Mizushima, S. (1997). The hydrophobic region of signal peptides is a determinant for SRP recognition and protein translocation across the ER membrane. *J Biochem (Tokyo)* 121:270–277.
8. Chirico, W.J., Waters, M.G., and Blobel, G. (1988). 70K heat shock related proteins stimulate protein translocation into microsomes. *Nature* 332:805–810.
9. McClellan, A.J., Endres, J.B., Vogel, J.P., Palazzi, D., Rose, M.D., and Brodsky, J.L. (1998). Specific molecular chaperone interactions and an ATP-dependent conformational change are required during posttranslational protein translocation into the yeast ER. *Mol Biol Cell* 9:3533–3545.
10. Plath, K., and Rapoport, T.A. (2000). Spontaneous release of cytosolic proteins from posttranslational substrates before their transport into the endoplasmic reticulum. *J Cell Biol* 151:167–178.
11. Jungnickel, B., and Rapoport, T.A. (1995). A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. *Cell* 82:261–270.
12. Belin, D., Bost, S., Vassalli, J.D., and Strub, K. (1996). A two-step recognition of signal sequences determines the translocation efficiency of proteins. *EMBO J* 15:468–478.
13. Crowley, K.S., Liao, S., Worrell, V.E., Reinhart, G.D., and Johnson, A.E. (1994). Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* 78:461–471.
14. Spiess, M. (1995). Heads or tails—what determines the orientation of proteins in the membrane. *FEBS Lett* 369:76–79.
15. Higy, M., Junne, T., and Spiess, M. (2004). Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* 43:12716–12722.
16. Shelness, G.S., Lin, L., and Nicchitta, C.V. (1993). Membrane topology and biogenesis of eukaryotic signal peptidase. *J Biol Chem* 268:5201–5208.
17. Lyman, S.K., and Schekman, R. (1997). Binding of secretory precursor polypeptides to a translocon subcomplex is regulated by BiP. *Cell* 88:85–96.

18. Plath, K., Mothes, W., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* 94:795–807.
19. Hann, B.C., and Walter, P. (1991). The signal recognition particle in *S. cerevisiae*. *Cell* 67:131–144.
20. Mutka, S.C., and Walter, P. (2001). Multifaceted physiological response allows yeast to adapt to the loss of the signal recognition particle-dependent protein-targeting pathway. *Mol Biol Cell* 12:577–588.
21. Lauring, B., Kreibich, G., and Weidmann, M. (1995). The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. *Proc Natl Acad Sci USA* 92:9435–9439.
22. Neuhof, A., Rolls, M.M., Jungnickel, B., Kalies, K.U., and Rapoport, T.A. (1998). Binding of signal recognition particle gives ribosome/nascent chain complexes a competitive advantage in endoplasmic reticulum membrane interaction. *Mol Biol Cell* 9:103–115.
23. Raden, D., and Gilmore, R. (1998). Signal recognition particle-dependent targeting of ribosomes to the rough endoplasmic reticulum in the absence and presence of the nascent polypeptide-associated complex. *Mol Biol Cell* 9:117–130.
24. Mothes, W., Jungnickel, B., Brunner, J., and Rapoport, T.A. (1998). Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane. *J Cell Biol* 142:355–364.
25. von Heijne, G. (1985). Signal sequences. The limits of variation. *J Mol Biol* 184:99–105.
26. von Heijne, G., and Manoil, C. (1990). Membrane proteins: From sequence to structure. *Protein Eng* 4:109–112.
27. von Heijne, G. (1990). The signal peptide. *J Membr Biol* 115:195–201.
28. von Heijne, G. (1989). Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* 341:456–458.
29. Hartmann, E., Rapoport, T.A., and Lodish, H.F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc Natl Acad Sci USA* 86:5786–5790.
30. Goder, V., and Spiess, M. (2001). Topogenesis of membrane proteins: determinants and dynamics. *FEBS Lett* 504:87–93.
31. Martoglio, B., and Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. *Trends Cell Biol* 8:410–415.
32. Hegde, R.S., and Bernstein, H.D. (2006). The surprising complexity of signal sequences. *Trends Biochem Sci* 31(10):563–571.
33. Pohlschroder, M., Prinz, W.A., Hartmann, E., and Beckwith, J. (1997). Protein translocation in the three domains of life: variations on a theme. *Cell* 91:563–566.
34. Jungnickel, B., Rapoport, T.A., and Hartmann, E. (1994). Protein translocation: common themes from bacteria to man. *FEBS Lett* 346:73–77.
35. Johnson, A.E. (2005). The co-translational folding and interactions of nascent protein chains: a new approach using fluorescence resonance energy transfer. *FEBS Lett* 579:916–920.
36. Crowley, K.S., Reinhart, G.D., and Johnson, A.E. (1993). The signal sequence moves through a ribosomal tunnel into a noncytoplasmic aqueous environment at the ER membrane early in translocation. *Cell* 73:1101–1115.
37. Wiedmann, M., Kurzchalia, T.V., Hartmann, E., and Rapoport, T.A. (1987). A signal sequence receptor in the endoplasmic reticulum membrane. *Nature* 328:830–833.
38. Kellaris, K.V., Bowen, S., and Gilmore, R. (1991). ER translocation intermediates are adjacent to a nonglycosylated 34-kD integral membrane protein. *J Cell Biol* 114:21–33.

39. High, S., Gorlich, D., Wiedmann, M., Rapoport, T.A., and Dobberstein, B. (1991). The identification of proteins in the proximity of signal-anchor sequences during their targeting to and insertion into the membrane of the ER. *J Cell Biol* 113:35–44.
40. Thrift, R.N., Andrews, D.W., Walter, P., and Johnson, A.E. (1991). A nascent membrane protein is located adjacent to ER membrane proteins throughout its integration and translation. *J Cell Biol* 112:809–821.
41. Krieg, U.C., Johnson, A.E., and Walter, P. (1989). Protein translocation across the endoplasmic reticulum membrane: identification by photocross-linking of a 39-kD integral membrane glycoprotein as part of a putative translocation tunnel. *J Cell Biol* 109:2033–2043.
42. Gorlich, D., Hartmann, E., Prehn, S., and Rapoport, T.A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* 357:47–52.
43. Hartmann, E., Gorlich, D., Kostka, S., Otto, A., Kraft, R., Knespel, S., Burger, E., Rapoport, T.A., and Prehn, S. (1993). A tetrameric complex of membrane proteins in the endoplasmic reticulum. *Eur J Biochem* 214:375–381.
44. Gorlich, D., and Rapoport, T.A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75:615–630.
45. Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T.A. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J Cell Biol* 134:25–35.
46. Fons, R.D., Bogert, B.A., and Hegde, R.S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J Cell Biol* 160:529–539.
47. Mothes, W., Prehn, S., and Rapoport, T.A. (1994). Systematic probing of the environment of a translocating secretory protein during translocation through the ER membrane. *EMBO J* 13:3973–3982.
48. High, S., Martoglio, B., Gorlich, D., Andersen, S.S.L., Ashford, A.J., Giner, A., Hartmann, E., Prehn, S., Rapoport, T.A., Dobberstein, B., and Bruner, J. (1993). Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. *J Biol Chem* 268:26745–26751.
49. Martoglio, B., Hofmann, M.W., Brunner, J., and Dobberstein, B. (1995). The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. *Cell* 81:207–214.
50. Sanders, S.L., Whitfield, K.M., Vogel, J.P., Rose, M.D., and Schekman, R.W. (1992). Sec61p and BiP directly facilitate polypeptide translocation into the ER. *Cell* 69:353–365.
51. Musch, A., Wiedmann, M., and Rapoport, T.A. (1992). Yeast Sec proteins interact with polypeptides traversing the endoplasmic reticulum membrane. *Cell* 69:343–352.
52. Heinrich, S.U., Mothes, W., Brunner, J., and Rapoport, T.A. (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* 102:233–244.
53. Mothes, W., Heinrich, S.U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J., and Rapoport, T.A. (1997). Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* 89:523–533.
54. Wang, L., Miller, A., Rusch, S.L., and Kendall, D.A. (2004). Demonstration of a specific *Escherichia coli* SecY-signal peptide interaction. *Biochemistry* 43:13185–13192.
55. Do, H., Falcone, D., Lin, J., Andrews, D.W., and Johnson, A.E. (1996). The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* 85:369–378.

56. Heinrich, S.U., and Rapoport, T.A. (2003). Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. *EMBO J* 22:3654–3663.
57. Van den Berg, B., Clemons, W.M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427:36–44.
58. Gorlich, D., Prehn, S., Hartmann, E., Kalies, K.U., and Rapoport, T.A. (1992). A mammalian homolog of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation. *Cell* 71:489–503.
59. Menetret, J.F., Hegde, R.S., Heinrich, S.U., Chandramouli, P., Ludtke, S.J., Rapoport, T.A., and Akey, C.W. (2005). Architecture of the ribosome-channel complex derived from native membranes. *J Mol Biol* 348:445–457.
60. Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T.A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81:561–570.
61. Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *EMBO J* 15:1482–1494.
62. Wittke, S., Dunnwald, M., Albertsen, M., and Johnsson, N. (2002). Recognition of a subset of signal sequences by Ssh1p, a Sec61p-related protein in the membrane of endoplasmic reticulum of yeast *Saccharomyces cerevisiae*. *Mol Biol Cell* 13:2223–2232.
63. Zimmermann, R. (1998). The role of molecular chaperones in protein transport into the mammalian endoplasmic reticulum. *Biol Chem* 379:275–282.
64. Hamman, B.D., Chen, J.C., Johnson, E.E., and Johnson, A.E. (1997). The aqueous pore through the translocon has a diameter of 40–60 A during cotranslational protein translocation at the ER membrane. *Cell* 89:535–544.
65. Hamman, B.D., Hendershot, L.M., and Johnson, A.E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. *Cell* 92:747–758.
66. Liao, S., Lin, J., Do, H., and Johnson, A.E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* 90:31–41.
67. Wirth, A., Jung, M., Bies, C., Frien, M., Tyedmers, J., Zimmermann, R., and Wagner, R. (2003). The Sec61p complex is a dynamic precursor activated channel. *Mol Cell* 12:261–268.
68. Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol Cell* 10:769–778.
69. Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997). Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* 278:2123–2126.
70. Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* 107:361–372.
71. Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., and Akey, C.W. (2000). The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 6:1219–1232.
72. Morgan, D.G., Menetret, J.F., Neuhof, A., Rapoport, T.A., and Akey, C.W. (2002). Structure of the mammalian ribosome-channel complex at 17A resolution. *J Mol Biol* 324:871–886.

73. Savitz, A.J., and Meyer, D.I. (1993). 180-kD ribosome receptor is essential for both ribosome binding and protein translocation. *J Cell Biol* 120:853–863.
74. Clemons, W.M., Jr., Menetret, J.F., Akey, C.W., and Rapoport, T.A. (2004). Structural insight into the protein translocation channel. *Curr Opin Struct Biol* 14:390–396.
75. Duong, F. (2003). Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. *EMBO J* 22:4375–4384.
76. Junne, T., Schwede, T., Goder, V., and Spiess, M. (2006). The plug domain of yeast sec61p is important for efficient protein translocation, but is not essential for cell viability. *Mol Biol Cell* 17:4063–4068.
77. Cannon, K.S., Or, E., Clemons, W.M., Jr., Shibata, Y., and Rapoport, T.A. (2005). Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *J Cell Biol* 169:219–225.
78. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., III, Ban, N., and Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* 438:318–324.
79. Ooi, C.E., and Weiss, J. (1992). Bidirectional movement of a nascent polypeptide across microsomal membranes reveals requirements for vectorial translocation of proteins. *Cell* 71:87–96.
80. Hegde, R.S., and Lingappa, V.R. (1996). Sequence-specific alteration of the ribosome-membrane junction exposes nascent secretory proteins to the cytosol. *Cell* 85:217–228.
81. Hegde, R.S., Voigt, S., Rapoport, T.A., and Lingappa, V.R. (1998). TRAM regulates the exposure of nascent secretory proteins to the cytosol during translocation into the endoplasmic reticulum. *Cell* 92:621–631.
82. Rutkowski, D.T., Lingappa, V.R., and Hegde, R.S. (2001). Substrate-specific regulation of the ribosome-translocon junction by N-terminal signal sequences. *Proc Natl Acad Sci USA* 98:7823–7828.
83. Beckmann, R., Spahn, C.M., Frank, J., and Blobel, G. (2001). The active 80S ribosome-Sec61 complex. *Cold Spring Harb Symp Quant Biol* 66:543–554.
84. Nicchitta, C.V., and Blobel, G. (1993). Lumenal proteins of the mammalian endoplasmic reticulum are required to complete protein translocation. *Cell* 73:989–998.
85. Tyedmers, J., Lerner, M., Wiedmann, M., Volkmer, J., and Zimmermann, R. (2003). Polypeptide-binding proteins mediate completion of co-translational protein translocation into the mammalian endoplasmic reticulum. *EMBO Rep* 4:505–510.
86. Dierks, T., Volkmer, J., Schlenstedt, G., Jung, C., Sandholzer, U., Zachmann, K., Schlotterhose, P., Neifer, K., Schmidt, B., and Zimmermann, R. (1996). A microsomal ATP-binding protein involved in efficient protein transport into the mammalian endoplasmic reticulum. *EMBO J* 15:6931–6942.
87. Klappa, P., Mayinger, P., Pipkorn, R., Zimmermann, M., and Zimmermann, R. (1991). A microsomal protein is involved in ATP-dependent transport of presecretory proteins into mammalian microsomes. *EMBO J* 10:2795–2803.
88. Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc Natl Acad Sci USA* 92:9643–9646.
89. Lyman, S.K., and Schekman, R. (1995). Interaction between BiP and Sec63p is required for the completion of protein translocation into the ER of *Saccharomyces cerevisiae*. *J Cell Biol* 131:1163–1171.
90. Matlack, K.E., Plath, K., Misselwitz, B., and Rapoport, T.A. (1997). Protein transport by purified yeast Sec complex and Kar2p without membranes. *Science* 277:938–941.

91. Corsi, A.K., and Schekman, R. (1997). The lumenal domain of Sec63p stimulates the ATPase activity of BiP and mediates BiP recruitment to the translocon in *Saccharomyces cerevisiae*. *J Cell Biol* 137:1483–1493.
92. Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* 97:553–564.
93. Veenendaal, A.K., van der Does, C., and Driessen, A.J. (2004). The protein-conducting channel SecYEG. *Biochim Biophys Acta* 1694:81–95.
94. Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* 116:725–736.
95. McCormick, P.J., Miao, Y., Shao, Y., Lin, J., and Johnson, A.E. (2003). Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. *Mol Cell* 12:329–341.
96. Haigh, N.G., and Johnson, A.E. (2002). A new role for BiP: Closing the aqueous translocon pore during protein integration into the ER membrane. *J Cell Biol* 156:261–270.
97. Alder, N.N., and Johnson, A.E. (2004). Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *J Biol Chem* 279:22787–22790.
98. Ota, K., Sakaguchi, M., von Heijne, G., Hamasaki, N., and Mihara, K. (1998). Forced transmembrane orientation of hydrophilic polypeptide segments in multispanning membrane proteins. *Mol Cell* 2:495–503.
99. Sakaguchi, M. (2002). Autonomous and heteronomous positioning of transmembrane segments in multispanning membrane protein. *Biochem Biophys Res Commun* 296:1–4.
100. Skach, W.R., and Lingappa, V.R. (1993). Amino-terminal assembly of human P-glycoprotein at the endoplasmic reticulum is directed by cooperative actions of two internal sequences. *J Biol Chem* 268:23552–23561.
101. Lu, Y., Turnbull, I.R., Bragin, A., Carveth, K., Verkman, A.S., and Skach, W.R. (2000). Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. *Mol Biol Cell* 11:2973–2985.
102. Rapoport, T.A., Goder, V., Heinrich, S.U., and Matlack, K.E. (2004). Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol* 14:568–575.
103. White, S.H., and von Heijne, G. (2004). The machinery of membrane protein assembly. *Curr Opin Struct Biol* 14:397–404.
104. Meacock, S.L., Lecomte, F.J., Crawshaw, S.G., and High, S. (2002). Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Mol Biol Cell* 13:4114–4129.
105. Sadlish, H., Pitonzo, D., Johnson, A.E., and Skach, W.R. (2005). Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat Struct Mol Biol* 12:870–878.
106. Ismail, N., Crawshaw, S.G., and High, S. (2006). Active and passive displacement of transmembrane domains both occur during opsin biogenesis at the Sec61 translocon. *J Cell Sci* 119:2826–2836.
107. Borel, A.C., and Simon, S.M. (1996). Biogenesis of polytopic membrane proteins: membrane segments of P-glycoprotein sequentially translocate to span the ER membrane. *Biochemistry* 35:10587–10594.
108. Saksena, S., Summers, M.D., Burks, J.K., Johnson, A.E., and Braunagel, S.C. (2006). Importin-alpha-16 is a translocon-associated protein involved in sorting membrane proteins to the nuclear envelope. *Nat Struct Mol Biol* 13:500–508.

109. Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 4:181–191.
110. Tsai, B., Ye, Y., and Rapoport, T.A. (2002). Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat Rev Mol Cell Biol* 3:246–255.
111. Hampton, R.Y. (2002). Proteolysis and sterol regulation. *Annu Rev Cell Dev Biol* 18:345–378.
112. Davidson, N.O., and Shelness, G.S. (2000). APOLIPOPROTEIN B: mRNA editing, lipoprotein assembly, and presecretory degradation. *Annu Rev Nutr* 20:169–193.
113. Oyadomari, S., Yun, C., Fisher, E.A., Kreglinger, N., Kreibich, G., Oyadomari, M., Harding, H.P., Goodman, A.G., Harant, H., Garrison, J.L., Taunton, J., and Katze, M.G. (2006). Cotranslocational degradation protects the stressed endoplasmic reticulum from protein overload. *Cell* 126:727–739.
114. Arvan, P., and Castle, D. (1998). Sorting and storage during secretory granule biogenesis: looking backward and looking forward. *Biochem J* 322(pt. 3):593–610.
115. Kornberg, R.D. (1996). RNA polymerase II transcription control. *Trends Biochem Sci* 21:325–326.
116. Novina, C.D., and Roy, A.L. (1996). Core promoters and transcriptional control. *Trends Genet* 12:351–355.
117. Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* 21:327–335.
118. Verrijzer, C.P., and Tjian, R. (1996). TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem Sci* 21:338–342.
119. Fickett, J.W., and Hatzigeorgiou, A.G. (1997). Eukaryotic promoter recognition. *Genome Res* 7:861–878.
120. Randall, L.L., and Hardy, S.J. (1989). Unity in function in the absence of consensus in sequence: role of leader peptides in export. *Science* 243:1156–1159.
121. Gruss, O.J., Feick, P., Frank, R., and Dobberstein, B. (1999). Phosphorylation of components of the ER translocation site. *Eur J Biochem* 260:785–793.
122. Ou, W.J., Thomas, D.Y., Bell, A.W., and Bergeron, J.J. (1992). Casein kinase II phosphorylation of signal sequence receptor alpha and the associated membrane chaperone calnexin. *J Biol Chem* 267:23789–23796.
123. Prehn, S., Herz, J., Hartmann, E., Kurzchalia, T.V., Frank, R., Roemisch, K., Dobberstein, B., and Rapoport, T.A. (1990). Structure and biosynthesis of the signal-sequence receptor. *Eur J Biochem* 188:439–445.
124. Noel, P., and Cartwright, I.L. (1994). A Sec62p-related component of the secretory protein translocon from *Drosophila* displays developmentally complex behavior. *EMBO J* 13:5253–5261.
125. Holthuis, J.C., van Riel, M.C., and Martens, G.J. (1995). Translocon-associated protein TRAP delta and a novel TRAP-like protein are coordinately expressed with pro-opiomelanocortin in *Xenopus* intermediate pituitary. *Biochem J* 312(1):205–213.
126. Garrison, J.L., Kunkel, E.J., Hegde, R.S., and Taunton, J. (2005). A substrate-specific inhibitor of protein translocation into the endoplasmic reticulum. *Nature* 436:285–289.
127. Besemer, J., Harant, H., Wang, S., Oberhauser, B., Marquardt, K., Foster, C.A., Schreiner, E.P., de Vries, J.E., Dascher-Nadel, C., and Lindley, I.J. (2005). Selective inhibition of cotranslational translocation of vascular cell adhesion molecule 1. *Nature* 436:290–293.
128. Levine, C.G., Mitra, D., Sharma, A., Smith, C.L., and Hegde, R.S. (2005). The efficiency of protein compartmentalization into the secretory pathway. *Mol Biol Cell* 16:279–291.

129. Kang, S.W., Rane, N.S., Kim, S.J., Garison, J.L., Tauton, J., and Hegde, R.S. (2006). Substrate-specific attenuation of protein translocation during acute ER stress defines a pathway of pre-emptive quality control. *Cell* 127:999–1013.
130. Shaffer, K.L., Sharma, A., Snapp, E.L., and Hegde, R.S. (2005). Regulation of protein compartmentalization expands the diversity of protein function. *Dev Cell* 9:545–554.
131. Hori, O., Miyazaki, M., Tamatani, T., Ozawa, K., Takano, K., Okabe, M., Ikawa, M., Hartmann, E., Mai, P., Stern, D.M., Kitao, Y., and Ogawa, S. (2006). Deletion of SERP1/RAMP4, a component of the endoplasmic reticulum (ER) translocation sites, leads to ER stress. *Mol Cell Biol* 26:4257–4267.
132. Davila, S., Furu, L., Gharavi, A.G., Tian, X., Onoe, T., Qian, Q., Li, A., Cai, Y., Kamath, P.S., King, B.F., Azurmendi, P.J., and Tahvanainen, P. (2004). Mutations in SEC63 cause autosomal dominant polycystic liver disease. *Nat Genet* 36:575–577.
133. Waanders, E., te Morsche, R.H., de Man, R.A., Jansen, J.B., and Drenth, J.P. (2006). Extensive mutational analysis of PRKCSH and SEC63 broadens the spectrum of polycystic liver disease. *Hum Mutat* 27:830.
134. Rutz, C., Renner, A., Alken, M., Schulz, K., Beyermann, M., Wiesner, B., Rosenthal, W., and Schulein, R. (2006). The corticotropin-releasing factor receptor type 2a contains an N-terminal pseudo signal peptide. *J Biol Chem* 281:24910–24921.
135. Kim, S.J., Mitra, D., Salerno, J.R., and Hegde, R.S. (2002). Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev Cell* 2:207–217.
136. Goder, V., Bieri, C., and Spiess, M. (1999). Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *J Cell Biol* 147:257–266.

This page intentionally left blank

# 10

# *The Role of BiP/Kar2p in the Translocation of Proteins Across the ER Membrane*

KUNIO NAKATSUKASA • JEFFREY L. BRODSKY

*Department of Biological Sciences*
*University of Pittsburgh*
*Pittsburgh*
*PA 15260, USA*

## I. Abstract

The endoplasmic reticulum (ER) is where proteins destined for secretory organelles—including the ER, Golgi, lysosome/vacuole, and plasma membrane—are first recognized by components that facilitate their maturation and intracellular trafficking. These secretory pathway residents are synthesized by ER-associated ribosomes and are then translocated or integrated into the ER membrane, a process that can occur either post-translationally or cotranslationally. A dedicated protein machine, known as the Sec61 complex, forms an aqueous channel in the ER membrane and thus facilitates the translocation reaction. Numerous studies have demonstrated that the 70-kDa heat-shock protein (Hsp70) in the ER lumen plays a critical role during protein translocation. This molecular chaperone not only drives the translocation of proteins across the ER membrane, but it helps maintain the permeability barrier of the ER during translocation and assists in the folding of the newly translocated protein. In this chapter, we will first describe the structure and function of Hsp70 chaperones and will then introduce the many components of the ER translocation

THE ENZYMES, Vol. XXV

ISSN NO: 1874-6047

DOI: 10.1016/S1874-6047(07)25010-3

machinery. Next, we will discuss how the lumenal Hsp70, known as BiP/Kar2p, catalyzes secreted protein translocation and folding in both the yeast and mammalian ER. Finally, we will review how BiP/Kar2p helps maintain ER homeostasis when aberrant proteins accumulate, and will raise important, unanswered questions in this field.

## II. Hsp70

The 70-kDa heat-shock protein (Hsp70) proteins are abundant chaperones that are present in eukaryotic cells, in eubacteria, and in many archaea [1, 2]. In eukaryotes, Hsp70s are found in the cytosol, chloroplasts, mitochondria, and in the endoplasmic reticulum (ER) lumen. Under stress conditions, such as heat shock, the expression of Hsp70s is induced to prevent aggregation and assist in the refolding of damaged proteins. However, Hsp70s also play important roles under normal conditions, and therefore many are constitutively expressed and are sometimes termed Hsc70s for heat-shock *cognate* proteins of ~70 kDa. Hsp70/Hsc70s assist in the folding of newly translated proteins, facilitate the translocation of proteins across biological membranes, disassemble oligomeric multiprotein complexes, and catalyze the proteolytic degradation of misfolded proteins. All of these activities rely on the adenosine 5′-triphosphate (ATP)-regulated association and disassociation of Hsp70s with hydrophobic polypeptide substrates [3–5]. However, the functional versatility of Hsp70s in the cell is achieved by two mechanisms. First, the Hsp70 genes have evolved to encode unique recognition and targeting motifs, which help diversify their function. Second, cochaperones target Hsp70 for unique roles in cellular activities. One class of cochaperones are known as Hsp40s and are distinguished by the presence of an ~70 amino acid ''J-domain''; these Hsp40s accelerate the ATPase activity of Hsp70s. Another class of cochaperones is Hsp70 nucleotide exchange factors (NEFs). The evolutionary diversity of Hsp70s is perhaps best understood for the yeast *Saccharomyces cerevisiae*. The genome of this yeast encodes ~14 Hsp70 proteins [6] and probably ~20 Hsp40 [7, 8], and different combinations of Hsp70–Hsp40 pairs function in processes as diverse as transcription, membrane fusion, and protein translation, translocation, and degradation [9].

### A. The Structure of Hsp70

Hsp70s consist of an ~44-kDa N-terminal nucleotide-binding domain (NBD) which binds and hydrolyzes ATP, an ~18-kDa substrate-binding domain (SBD), and an ~10-kDa C-terminal variable domain [10, 11]

(Figure 10.1A). Some nonorganellar eukaryotic Hsp70s have a conserved EEVD sequence at the extreme C-terminus that binds tetratricopeptide repeat (TPR) proteins, which further diversify Hsp70 function and link this chaperone to other chaperone systems [12].

Some time ago, high-resolution structures were solved by X-ray crystallography for the ATPase domain of bovine Hsc70 [10] and the SBD of an *Escherichia coli* Hsp70, DnaK [11]. The ATPase domain of bovine Hsc70 is composed of two subdomains, separated by a cleft that contains the nucleotide-binding site [10]. The nucleotide is bound at the base of the cleft together with one $Mg^{2+}$ and two $K^{+}$ ions and is almost entirely solvent-inaccessible. The SBD of DnaK has an 8-stranded antiparallel β-sandwich, and the peptide is bound in an extended conformation in a channel formed by loops that connect the β-sheets and create a hydrophobic pocket. A distal α-helical domain covers this pocket, but does not contact the peptide. It has been suggested that this helical domain acts as a ''lid'' to stabilize peptide complexes and to control the kinetics of peptide binding and release [13, 14].

Although the ATPase domains with bound NEFs have been cocrystallized and their structures solved [15, 16], only recently was a structure of a nearly full-length Hsp70 solved by X-ray crystallography [17]. The structure suggests that the NBD and the SBD interact via an SBD-resident helix that inserts between two subdomains in the NBD, and the biochemical analysis of site-directed mutants supports many of the conclusions at which the authors of this groundbreaking study arrived. However, it is important to note that the structure was determined in the absence of bound

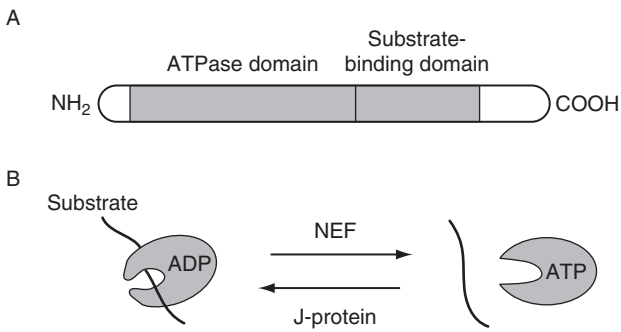



FIG. 10.1. (A) Schematic structure of the Hsp70 chaperone. (B) Catalytic cycle of Hsp70. Hsp70 cycles between the ADP and ATP-bound state as substrates cycle on and off. A J-domain-containing protein (J-protein), such as an Hsp40, accelerates the ATPase activity of Hsp70 by binding to the ATPase domain. A ''NEF'' may also enhance the ATPase activity of Hsp70 by increasing the rate of nucleotide exchange.

nucleotide—which may impact interdomain communication between the NBD and SBD—and in the absence of a peptide substrate; indeed, the SBD pocket in this structure is occupied by a disordered helix that resides elsewhere in the SBD. In contrast, NMR chemical shift analysis of the SBD of DnaK indicated that this domain exists in a high peptide-affinity state and that ATP must be bound for the SBD and NBD to interact [18], even though others have found that either adenosine 5′-diphosphate (ADP) or ATP supports NBD–SBD association [19]. More recent work indicates that the interdomain linker between the NBD and SBD plays a major role in mediating the ATP/ADP-dependent conformational change [19a]. Clearly, much work needs to be directed toward solving the structure of full-length Hsp70—in the presence of different nucleotides and with bound peptides and cochaperones—by both NMR and X-ray crystallography, and this goal represents a major challenge to those in the field.

### B. The Hsp70 Reaction Cycle

The Hsp70 catalytic cycle has been best defined for DnaK in the presence and absence of its Hsp40 cochaperones, DnaJ, and a NEF, GrpE (Figure 10.1B). In DnaK-assisted folding reactions, substrates repeatedly cycle on and off the chaperone as DnaK cycles between the ADP and ATP-bound state [20, 21]. The ATP-bound form of DnaK has a low affinity and fast substrate exchange rate because the α-helical latch over the peptide-binding cleft is open; in contrast, the ADP-bound form has a high affinity and slow exchange rates for the substrates because the α-helical latch is closed [22–25] (and see above). DnaK and other Hsp70s have a low basal ATPase activity, which is probably one reason that cochaperones, such as Hsp40 and NEFs, are essential for many Hsp70-catalyzed reactions. Indeed, DnaJ significantly accelerates DnaK's ATPase activity by contacting the underside of the Hsp70 NBD with the J-domain [13, 26–29]. As a result, a peptide substrate becomes trapped in the SBD of DnaK. In some cases, DnaJ, which also binds hydrophobic peptides [30], can deliver substrates to Hsp70 [20, 31–35]. In either case, the GrpE NEF helps free the bound substrate by increasing the rate of nucleotide exchange, and the addition of GrpE to DnaJ and DnaK synergistically enhances DnaK's ATPase activity [36]. As a result of the DnaK ATP/ADP cycle, peptides with overall hydrophobic character are continuously routed on and off the chaperone, which is critical for the myriad of Hsp70-catalyzed cellular functions. It is important to note, however, that Hsp70 does not itself significantly change the conformation of the bound substrate—in kinetic terms, it only decreases the frequency at which a substrate will occupy off-pathway intermediates.

### C. BiP/Kar2p

BiP, an abundant Hsp70 in the ER lumen, was originally identified both as an immunoglobulin heavy chain-*b*inding *p*rotein [37] and as a *g*lucose-*r*egulated *p*rotein, Grp78, because it accumulates following glucose starvation [38]. The N-terminal NBDs of BiP and a cytosolic Hsc70 are 66% identical and 81% similar [39], and the C-terminus contains an ER retention sequence [40, 41]. BiP is expressed in all cell lineages ranging from yeast to man, and interestingly yeast BiP was identified because it was the second gene isolated that is required for nuclear/ER membrane fusion (karyogamy) during the mating of **a** and α yeast cells, and was therefore called Kar2p [42–44]. Kar2p is ~50% similar to *E. coli* DnaK and the expression of mammalian BiP partially replaces Kar2p function in yeast [42, 45]. Because there is no correlation between defects in nuclear fusion and translocation in yeast containing various *kar2* mutant alleles [46], it is likely that Kar2p is directly involved in nuclear/ER membrane fusion, although its mechanism of action during this process remains mysterious.

Mammalian BiP and Kar2p in yeast, like all Hsp70s, bind and hydrolyze ATP, which in turn regulates peptide association/disassociation. And, consistent with the fact that Hsp40s help dictate Hsp70 function (see above), the yeast ER houses three DnaJ homologues that specify BiP/Kar2p action [8]: Sec63p [47], Scj1p [48], and Jem1p [49]. Sec63p is a multispanning membrane protein with an ER-disposed J-domain [47]. Scj1p and Jem1p are soluble ER lumenal proteins but Jem1p may peripherally associate with the inner ER membrane [50]. Sec63p catalyzes polypeptide translocation by positioning Kar2p at the ER and by stimulating its ATPase activity [51–54], whereas Scj1p is involved in secretory protein folding in the ER [48] and Jem1p is required for membrane fusion during karyogamy [49]. Other unique functions of these Kar2p cochaperones have been uncovered during the disposal of unfolded proteins in the ER [55] (and see below).

In the mammalian ER, five DnaJ homologues have been identified: ERdj1/ERj1/Mtj1 [56–60], ERdj2/mSec63 [61, 62], ERdj3/HEDJ/ERj3/ABBBP-2 [63–65], ERdj4/Mdg1 [66, 67], and ERdj5/JPDI [68, 69], all of which have been shown to stimulate the ATPase activity of BiP *in vitro*. As might be expected, the homologues also exhibit some functional diversity in the cell: ERj1p is a membrane protein but the N-terminal cytosolic domain binds to the 28S ribosomal RNA and contributes to the formation of a ribosome–translocon complex [70]. It has been proposed that ERj1p inhibits translation in the absence of BiP and regulates gene expression by interacting with the nuclear import factor, importinβ [71]. ERdj2/mSec63 is a mammalian homologue of yeast Sec63p and a small fraction

associates with the Sec61 protein [62], which forms the translocation channel (see below). ERdj3, 4, and 5 are induced when the ER is under stress, a condition that leads to new chaperone synthesis and degradation of aberrant proteins that may induce stress. As might be expected, then, ERdj3p directly associates with a number of unfolded or partially folded proteins in the ER along with BiP, thereby contributing to the folding process [72], and accumulating evidence suggests that ERdj4/Mdg1 and ERdj5/JPDI facilitate BiP-mediated protein folding in the ER [66–69]. Although GrpE homologues have not been identified in the yeast or mammalian ER, other classes of NEFs have been found in this compartment [72a]: one class is defined by BAP in mammals and by Sls1p/Sil1p in yeast [73, 74]. BAP and Sls1p/Sil1p bind to the ATPase domain of BiP and Kar2p, respectively. Another class is defined by Grp170 (in mammals) and Lhs1 (in yeast) [75, 76], which are employed to enhance BiP's ability to catalyze unique reactions in the ER (also see below).

Several reports suggest that BiP's activity can be regulated in other manners, which may further diversify its function in the cell. First, mammalian BiP can be isolated in monomeric and oligomeric forms, and it has been proposed that dimeric BiP can be activated upon conversion into the monomeric species when a peptide binds [77]. Second, BiP is posttranslationally modified by phosphorylation and ADP ribosylation. These modifications may be important to regulate the synthesis and polypeptide-binding activity of the chaperone [78]. Third, BiP is an important calcium-binding protein in the ER and in turn calcium may influence BiP's activity [79, 80]. Fourth, BiP expression is induced by ER stress [81, 82], which provides a feedback mechanism to temper subsequent insults to the ER. Although the induction of BiP by ER stress—particularly that brought about by the accumulation of unfolded proteins (i.e., the unfolded protein response or UPR [83])—clearly plays an important role in ER physiology, it is less obvious at present whether these other mechanisms of BiP regulation occupy an equally important position *in vivo*.

## III. Protein Translocation into the ER

In eukaryotes, the synthesis of most membrane proteins and soluble, secreted proteins begins on cytoplasmic ribosomes, which are then targeted to the cytoplasmic face of the ER membrane. These nascent polypeptides are subsequently integrated into the lipid bilayer of the ER membrane or are transported into the ER lumen, respectively. In this scenario, protein insertion or ''translocation'' into the ER most likely occurs during polypeptide synthesis, and is termed cotranslational translocation

(Figure 10.2A). In contrast, some soluble, secreted polypeptides can be synthesized exclusively in the cytoplasm and are translocated posttranslationally (Figure 10.2B); under these conditions, translation and translocation are uncoupled, and cytoplasmic chaperones are required to prevent the nascent, secreted polypeptides from aggregating [84, 85]. However, during cotranslational translocation, nascent polypeptide aggregation cannot occur in the cytoplasm because translation and translocation are coupled at the ER membrane; this reaction requires the actions of both the signal recognition particle (SRP) and the SRP receptor (SR) [86]. In yeast, ER-targeted proteins fall into three classes: SRP dependent, SRP independent, and those that are able to follow either pathway. These classes are distinguished by differences in the composition of their cleavable, N-terminal signal sequences [87], but it is not as obvious why a given polypeptide might have evolved to utilize one route versus another. Regardless of the pathway that is taken, the Sec61 complex forms the protein-conducting channel, and the multispanning Sec61 protein is the major component of the aqueous conduit into the ER [88–94].

## A. The Sec Complex

In mammals, the Sec61 complex is composed of Sec61$\alpha$, Sec61$\beta$, and Sec61$\gamma$, and in the yeast *S. cerevisiae*, the ER membrane contains homologous core components: Sec61p, Sbh1p, and Sss1p, respectively. Sec61$\alpha$/Sec61p spans the membrane ten times, with both N- and C-termini in the cytosol [95]. The second subunit, Sec61$\beta$/Sbh1p, spans the membrane once with the N-terminus in the cytosol. The third subunit, Sec61$\gamma$/Sss1p, is probably a single spanning protein with its N-terminus in the cytosol. In *S. cerevisiae*, there is a second Sec61-like complex (the Ssh1 complex) that is not essential for cell growth, but because it associates with ribosomes it may be involved exclusively in the cotranslational translocation [96].

After the targeting of a nascent polypeptide to the ER membrane, several proteins are required for translocation that function upstream of the Sec61 complex. For example, in *S. cerevisiae* the tetrameric Sec63p complex contains BiP/Kar2p, Sec63p, Sec71p, and Sec72p, and can associate with a larger complex that contains the Sec61 complex as well as Sec62p [53, 94, 97] (Figure 10.2B). Sec63p spans the membrane three times with the N-terminus in the lumen and the C-terminus in the cytosol, and as noted above deposits a J-domain in the ER lumen in order to recruit Kar2p [47] (Figure 10.2B). The extreme C-terminal 52 residues of Sec63p are predominantly acidic, and the last 14 of these have been shown to interact directly with Sec62p [98, 99]. Preceding this acidic domain, there is a so-called Brl domain (Brr2p-like) that is homologous to a repetitive sequence present in

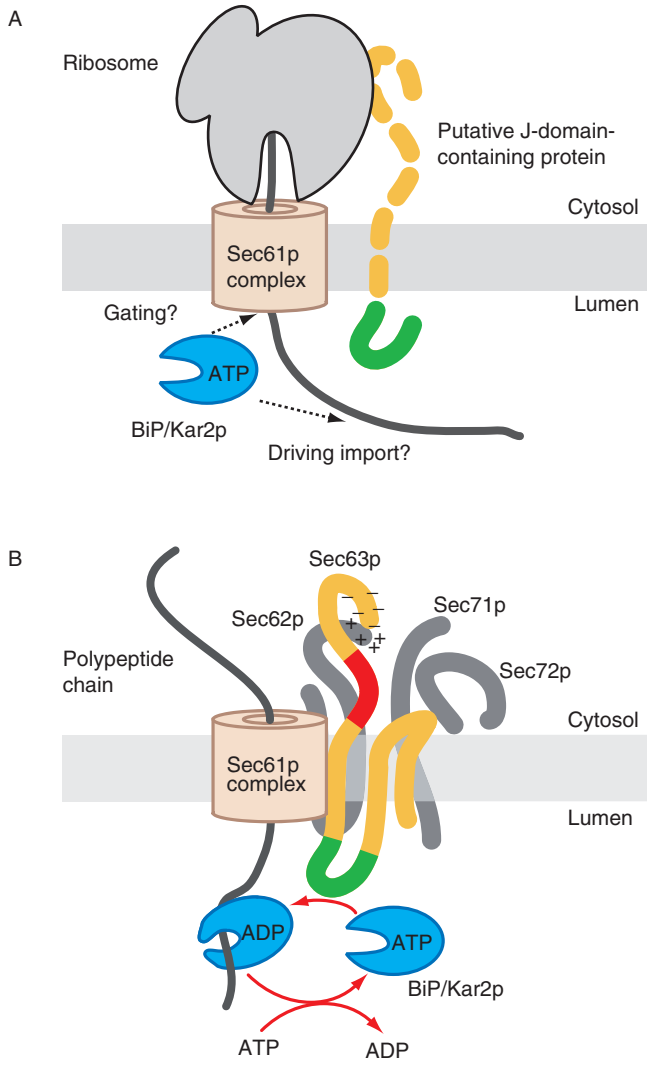



FIG. 10.2. (A) During cotranslational translocation, translation and translocation are coupled. Polypeptide chains are synthesized by the ER-bound ribosome and are transported into the ER lumen through the Sec61p complex, which includes Sec61p, Sbh1p, and Sss1p in yeast. The ER lumenal chaperone, BiP/Kar2p, may facilitate import by binding to the substrate. BiP/Kar2p may also seal the lumenal end of the translocon to prevent the passage of small molecules. Some J-domain-containing proteins (the J-domain is indicated in green) are proposed to function in these processes (see text for details). (B) During posttranslational translocation, translation and translocation are uncoupled. The ATPase activity of BiP/Kar2p is stimulated by the J-domain in Sec63p (the J-domain is indicated in green). BiP/Kar2p binds

Brr2p and other related members of the U5 200-kDa family of RNA helicases [100, 101].

Homologues of yeast Sec62p and Sec63p were identified in humans [61, 102] and in dog pancreas microsomes and interact with one another [62]. Even though mammalian Sec62p and Sec63p are abundant, only a minor fraction of these proteins associate with the Sec61 complex [62] and are found exclusively in ribosome-free complexes [102]. This stoichiometry is radically different from that in yeast, where it was estimated that essentially every Sec62p and Sec63p molecule associate with the Sec61-containing complex [97].

The assembly of the Sec complex has been analyzed in yeast and led to the identification of four distinct Sec61-containing complexes of 140, 280, 350, and 380 kDa [103]. The two smaller complexes may correspond respectively to a dimer and tetramer of the Sec61 complex consisting of Sec61p (53 kDa), Sbh1p (8.7 kDa), and Sss1p (8.9 kDa), which forms the core of the ER translocation channel and is similar to the 140-kDa complex identified in the mammalian ER [91]. The largest complex contains Sec61p, Sbh1p, Sss1p, Sec63p, Sec71p, Sec72p, and Sec62p [97] (Figure 10.2B) and the second largest 350-kDa complex contains all of these components except Sec62p. A central member of the largest complexes, which is consistent with its interaction with BiP/Kar2p and the Sec61p complex, is Sec63p. The Sec63p Brl domain is required to assemble the largest two complexes and its deletion causes both a post- and cotranslational translocation defect, but the acidic domain is required to form only the 380 kDa complex and its deletion causes a posttranslational translocation defect. It was therefore suggested that the 380 kDa complex is specialized for posttranslational translocation (which requires Sec62, consistent with previous data [104]) whereas the 350 kDa complex is required for cotranslational translocation [103].

Structural studies have suggested that the ER translocon is an oligomer that contains 3–4 copies of the Sec61 complex [92, 93, 105, 106]. The homologous translocon in *E. coli* contains two or four copies of the conserved SecYEG complex [107, 108], and it was suggested that the pore may be located at the interface between several copies of Sec61/SecY heterotrimer. Consistent with the large aperture that would derive from placement of the pore at this interface, Johnson and colleagues used fluorescence

to the polypeptide chain emerging from the Sec61p complex and facilitates translocation. The acidic domain at the C-terminus of Sec63p interacts with the C-terminus of Sec62p. The Brl domain in Sec63p (red region) is required to assemble a multiprotein complex that along with BiP is sufficient for translocation *in vitro* [97] and that contains the Sec61p, Sec62p, Sec63p, Sec71p, and Sec72p complexes [103].

measurements of an actively translocating polypeptide and estimated the pore diameter in the mammalian ER to be 40–60 Å [109]. However, a high-resolution X-ray crystal structure of the SecY complex from a single-celled archaeon, *Methanococcus jannaschii*, suggested instead that the functional pore may be formed by a single Sec61/SecY heterotrimer [110]. The role of oligomerization, which was observed in the previous studies, is not known, but heterotrimer fusion might be necessary to form a larger channel under certain circumstances or to accommodate the translocation of specific substrates. It is also possible that the mammalian and bacterial translocation channels function distinctly, and/or that the unique technique employed cannot allow direct conclusions to be made at this time. The solution to this discrepancy may finally emerge from visualizing the structure of the complex with an embedded, translocating polypeptide.

## B. Posttranslational Translocation

### *1. Substrate Targeting to the Sec Complex*

As noted above, eukaryotes can translocate some proteins across the ER membrane posttranslationally. This pathway may be largely restricted to lower eukaryotes such as yeast *S. cerevisiae* because in mammals only substrates <70 amino acids are translocated posttranslationally [111] and yeast can survive in the absence of SRP, which is required for cotranslational translocation [112, 113] (see below). To prevent aggregation, these fully synthesized polypeptides interact with cytosolic chaperones such as Hsp70 (Ssa1p in yeast), Hsp40 (Ydj1p in yeast), and the TRiC/CCT Hsp60-like chaperone complex, which together may keep the precursors in a soluble, translocation-competent state [114–116]. After the bound cytosolic chaperones are released, possibly because they spontaneously dissociate [116] the precursor proteins are recognized by the Sec complex [117]. Photo- and chemical cross-linking experiments revealed that the N-terminal signal sequence simultaneously contacts Sec61p and Sec62p, whereas the C-terminal portion of the posttranslationally translocated substrate contacts Sec62p, Sec71p, and Sec72p [118–120]. This association occurs in the absence of Kar2p function [117, 119, 120] but requires the association of the other membrane-resident factors [121]. Although the driving force for the initial movement of a posttranslationally translocated substrate into the ER is unknown, the polypeptide chain inserts into the channel as a hairpin and exposes a small segment to the ER lumen. Subsequent movement of the polypeptide through the channel requires BiP/Kar2p [14, 51, 118, 122].

### 2. *How Does BiP/Kar2p Contribute to the Unidirectional Movement of a Polypeptide During Posttranslational Translocation?*

How does BiP/Kar2p facilitate polypeptide translocation into the ER? One mechanism has been termed the ''Brownian ratchet'' [123, 124]. In this model, a protein in the ''translocon'' (i.e., a generic translocation channel) is free to oscillate passively in either direction due to Brownian motion/thermal energy. When a sufficiently long segment of the polypeptide chain extends into the ER lumen, BiP/Kar2p captures the substrate and prevents backsliding. As subsequently longer segments of the polypeptide extend into the ER and BiP continues to bind in increasing amounts, unidirectional translocation is favored. Mathematical modeling suggests that known translocation rates can be described by a Brownian ratchet mechanism [125], and that a Brownian ratchet could be used to drive translocation by increasing the entropy of the polypeptide at the membrane; that is, the chaperone would prevent the translocating polypeptide from associating with the inner surface of the bilayer, which would otherwise minimize its diffusion potential [126]. This model also predicts that the purpose of Sec63p, the J-domain-containing partner for BiP, is to localize the chaperone near the translocation site, and that ATP hydrolysis is required to lock BiP onto the substrate.

This model was experimentally tested using purified Sec complex and prepro-α-factor—a small yeast mating prepheromone that posttranslationally translocates—as a substrate [51]. Consistent with the Brownian ratchet model, multiple BiP molecules bound the substrate during its translocation through the channel and the interaction between BiP and the Sec63p J-domain of Sec63p and subsequent ATP hydrolysis were required to trap the substrate [52]. Indeed, a partially translocated substrate, which was stalled by appending a bulky group onto the C-terminus of prepro-α-factor, slid backward if ATP was removed. The model was further supported by the observed translocation of prepro-α-factor into lumenally depleted ER microsomes that only contained antibodies against prepro-α-factor [51].

Another model, which may be complementary to the Brownian ratchet, has been termed the ''translocation motor.'' In this model, BiP/Kar2p acts as a force-generating motor. The motor model was first proposed to explain how a folded, mitochondrial-targeted protein is unfolded during its import [127, 128]; these two events must be coupled because folded precursor proteins are too bulky to pass through most protein translocation channels [129, 130]. Because the spontaneous unfolding of many proteins may be extremely slow, a strong force must be applied to pull a polypeptide into the mitochondria as a folded domain unravels. Support emerged for this model from the fact that mitochondria containing an *hsp70* mutant allele

cannot import tightly folded precursor proteins, but can import loosely folded proteins or proteins that had been artificially unfolded [131]. In this model, one must assume that the Hsp70 is attached simultaneously both to the precursor and to the translocation channel (or an ER membrane J-domain protein, such as Sec63p), and that ATP hydrolysis induces a conformational change in the Hsp70 that drives polypeptide import using the J-domain as a fulcrum. It is important to note that the translocation of unfolded domains may also require a motor if a polypeptide interacts favorably with the inner wall of the channel [132]. In fact, one mathematical model suggested that the Brownian ratchet provides insufficient force to support posttranslational translocation because of these substrate–translocon contacts [133]. It is also worth noting that translocation is inhibited in yeast containing Kar2p/BiP mutants that are unable to exhibit an ATP-dependent conformational change but that still bind polypeptide substrates [134], lending some support to this model.

### *3. Other BiP-Interacting Factors May Be Required During Posttranslational Translocation into the ER*

In addition to Kar2p and Sec63p, the Lhs1p (the yeast homologue of Grp170) and Sls1p/Sil1p NEFs were shown to contribute to the translocation of some proteins into the yeast ER [135]. Lhs1p and Kar2p coordinately regulate one another's ATPase activity [76], although cells depleted for Lhs1p show defects only in posttranslational translocation [136]. Similarly, the deletion of *SLS1*/*SIL1* does not affect translocation, but cells containing both a conditional *lhs1* allele and a *kar2* mutant allele (in the ATPase domain) exhibit severe post- and cotranslational translocation defects [74, 135]. These data suggest that the NEFs are required for maximal translocation efficiency, although it is possible that Sls1p/Sil1p cooperates with Kar2p during the folding of nascent polypeptides in the ER, and that defects in this process back up the import machinery. It is also important to consider that these mutations induce ER stress, and the subsequent upregulation of one of many UPR-inducible factors (see above) may have secondary consequences on translocation. Regardless, it will be exciting in the future to better define the functions of these newly characterized NEFs in the secretory pathway.

## C. Cotranslational Translocation

### *1. Nascent Polypeptide Targeting in the Cotranslational Translocation Pathway*

During cotranslational translocation, the signal sequence emerges from the ribosome and is immediately recognized by the SRP. Mammalian SRP is a complex of a 7S RNA molecule and six proteins, SRP9, SRP14, SRP19,

SRP54, SRP68, and SRP72 [86, 137]. SRP slows or temporarily halts continued polypeptide elongation and then targets the ribosome/nascent polypeptide complex to the Sec61 complex in the ER membrane via its interaction with the SR. SR is a heterodimer of the SRα and SRβ proteins that are peripheral and integral membrane GTPases, respectively. A GTP-dependent interaction of SRP with SR triggers a series of event that includes the release of the signal sequence from SRP and the release of SR from the ribosome and the translocation machinery, respectively. After SRP release, protein synthesis resumes and the nascent chain migrates into the translocation pore [138–143].

### 2. *How Does BiP/Kar2p Contribute to the Cotranslational Translocation?*

During posttranslational translocation in yeast, BiP/Kar2p binding to the incoming polypeptide chain is necessary for the unidirectional movement of the polypeptide, at least in *in vitro* reconstituted systems [51]. During cotranslational translocation, one might envision that BiP would be dispensable because the ribosome—as it synthesizes an elongating polypeptide chain—may ''push'' a polypeptide into the ER. Consistent with this hypothesis, Görlich and Rapoport [91] reported that cotranslational protein translocation can occur in reconstituted proteoliposomes that lack lumenal factors. In addition, it was shown that the translocation of a long N-terminal domain that is mediated by type I signal-anchor sequence does not require BiP [144].

In contrast, other publications have supported the requirement for lumenal factors—including BiP/Kar2p—during the cotranslational translocation reaction. First, the ER lumenal requirements during cotranslational translocation were analyzed by Nicchitta and Blobel [145], and these investigators discovered that lumenal proteins were necessary to translocate preproteins into mammalian microsomes. Of these ER lumenal proteins, ATP-binding proteins were shown to be a part of the protein ''translocase'' and were needed to facilitate presecretory protein insertion into Sec61 [146, 147]. Using proteoliposomes made from microsomal detergent extracts, the requisite ATP-binding proteins were identified as BiP (Grp78) and Grp170 (a homologue of yeast Lhs1p, see above) [148]. The BiP requirement was then confirmed using proteoliposomes prepared from microsomal detergent extracts that contained trapped reticuloplasm [149] and through the use of high-pH-washed vesicles in which individual lumenal factors could be inserted [150]. In fact, when the translocation of biotinylated precursor proteins was examined, BiP substituted for avidin in the lumen and avidin improved the translocation efficiency substantially [149]. These results suggested that BiP and/or other ER lumenal chaperones

directly facilitate cotranslocational translocation. Second, yeast cells containing a *kar2* mutant allele or that were depleted for Sec63p exhibited a specific cotranslational translocation defect [151, 152]. And more directly, ER-derived microsomes prepared from several *kar2* mutants and a *sec63* mutant that abrogates its ability to associate with BiP [53] were defective for the cotranslational translocation of a yeast preprotein [153].

What is the resolution to these conflicting data? Although simplified reconstituted systems suggested that BiP/Kar2p is not essential for import, the lumenal chaperones might greatly increase the translocation efficiency, and depending on the laboratory and conditions employed, BiP might either be required or simply act as a facilitator. Another possibility is that reactions that start off cotranslationally convert to the posttranslational mode as translocation proceeds. Depending on the conditions and/or substrate, a BiP requirement might not be evident. It is thus hoped that with the continued development of new substrates and *in vitro* systems (e.g., the improved cotranslational translocation system developed by Stirling and coworkers [152]) these discrepancies may be resolved.

### *3. BiP May Maintain the Permeability Barrier During Translocation*

During cotranslational translocation in mammals, the ER translocon is dynamically resized. Using fluorescently labeled, translocating polypeptides and by artificially introducing fluorescent quenchers into the ER, it was shown that the channel expands from an inner diameter of 9–15 Å in the ribosome-free state [150] to a diameter of 40–60 Å in the active, ribosome-bound state [109, 154]. Even given this large size, the translocation channel should prevent the passage of small molecules during translocation in order to maintain the lumenal ionic and oxidizing environments. In fact, Johnson and coworkers demonstrated that the formation of a sealed ribosome–translocon junction prevents the passage of small molecules [155] and that the lumenal side of the ribosome-free translocon pore is sealed either directly or indirectly by BiP [150]. The proposed mechanism for this BiP-mediated pore closure and opening is similar to the normally employed Hsp70 chaperone cycle: BiP seals the lumenal end of the translocon in its ADP-bound form via an interaction both between the SBD and a translocon-associated component, and between the NBD and a J-domain-containing protein. After the SRP-dependent targeting of a ribosome nascent chain complex to the ER, the BiP-containing gate continues to seal the translocon from the lumenal side, but after the nascent chain reaches a length of $\sim$70 residues [155], BiP is released from the translocon on exchange of ADP for ATP. Once the translation of the substrate is complete, ribosome-free translocons are resealed by BiP [156]. The most

likely candidates for the J-domain containing protein in this cycle are mSec63/ERdj2 and/or ERdj1/ERj1/Mtj1 [71]. Although it is not yet known whether this BiP-mediated sealing of the translocon is essential for the translocation reaction itself, this model is consistent with the cotranslational translocation defects observed in *kar2* and *sec63* mutant yeast [151, 153] (see above). Alternatively, the permeability barrier may be maintained by a constraint of the translocon pore that is achieved by a movement of a short helix (called the ''plug'') to the center of the archael pore that was visualized by X-ray crystallography [110, 157] (see above). However, the function of the plug in maintaining the permeability barrier is still being debated, and as noted, high-resolution structures of actively translocating translocation complexes will significantly help define how—or even if—the permeability barrier is maintained [158].

## IV. Folding of Nascent Proteins in the ER and ER-Associated Degradation (ERAD)

Proteins translocating into the ER immediately begin to fold and become posttranslationally modified. These modifications include signal sequence cleavage, the transfer and trimming of N-linked glycans, disulfide bond formation, and prolyl isomerization, and these events and protein folding extensively rely on the assistance of molecular chaperones and other factors in the ER lumen. Interactions with such ER residents and posttranslational processing typically begin as growing nascent chains enter the lumen through the pore [159–161], whereas some proteins may be posttranslationally folded to acquire their proper native conformation [162]. Regardless of when folding occurs, BiP plays an active role in this process by preventing the aggregation of exposed hydrophobic domains, thus retaining the nascent polypeptide on the folding pathway. In the mammalian ER, it is likely that BiP catalyzes folding as a component of a multimeric complex that includes Grp94 (an Hsp90 homologue in the ER), Grp170 (the Lhs1p homologue), protein disulfide isomerase (PDI), ERdj3, cyclophilin B (an FKBP immunophillin), ERp72 (a putative disulfide isomerase), UDP-glucosyltransferase, and SDF2-L1 (a homologue of *O*-mannosyl transferase) [163]. In yeast, the existence of an analogous folding machine has not been reported, although the J-domain-containing Scj1 protein participates in the productive folding of newly translocated proteins [48], assisting BiP/Kar2p in this critical task [164].

The ER also contains a lectin chaperone system, which has been extensively analyzed in mammals. Soon after the N-linked, core oligosaccharide has been added to the nascent polypeptide in the ER, the two outermost

glucose residues are rapidly trimmed and the resulting glycan is recognized by the ER lectins calnexin and/or calreticulin. These chaperones prevent deleterious associations with other factors and prevent protein aggregation; they also recruit, ERp57, a disulfide isomerase that further catalyzes folding. Proteins are released from calnexin/calreticulin after cleavage of the innermost glucose residue by glucosidase II. If, however, the protein remains in a nonnative state, the UDP-glucose:glycoprotein glucosyltransferase (UGGT) recognizes the exposed misfolded domain and reglucosylates the N-linked glycan, thus triggering calnexin/calreticulin reassociation with the substrate [165–170]. This cycle ensures that nascent, secreted proteins are retained in the ER until they fold and thus escape the action of the UGGT and lectin binding.

During translocation, a choice is made between the BiP and lectin chaperone systems. Which chaperone complex is recruited probably depends on whether there is an N-linked glycan in the first 50 amino acids of the substrate. If a core oligosaccharide is added, BiP quickly hands the substrate off to calnexin and then perhaps to calreticulin. If instead a glycan modification is not encountered, BiP binds directly to the substrate [171, 172]. This sequential action of the chaperones is consistent with the placement of BiP at the lumenal face of the translocon: BiP is released from the translocon when the nascent chain reaches the length of ~70 amino acids (~40 inside the ribosome channel and ~30 in the translocon), the point at which residues in the nascent chain first ''see'' the lumenal space and can bind to BiP [155].

Although there is an elaborate system of chaperones that facilitates productive folding, some proteins cannot achieve their correct conformation because of a genetic error, environmental stress, stochastic misfolding, or changes in ER homeostasis (e.g., altered calcium levels, a change in the oxidative state or pH, or a decreased availability of sugars for N-linked glycosylation). These misfolded proteins are recognized as being defective and are degraded through a process known as ER-associated degradation (ERAD) [173–177]. Surprisingly, the degradation of the aberrant species does not occur in the ER but instead takes place after the misfolded protein has been retrotranslocated back to the cytoplasm. In the cytoplasm most ERAD substrates are polyubiquitinated and are then delivered to the proteasome, a multicatalytic protease. The channel required to retrotranslocate an ERAD substrate remains controversial; in some instances the Sec61 channel may be used, but ERAD-specific channels such as Derlin 1 [178, 179] have also been proposed.

Because of its central role in protein translocation and protein folding in the ER, it should come as no surprise that BiP/Kar2p is also required for ERAD. By analogy to the mechanism by which cytosolic Hsp70 and Hsp40

chaperones maintain substrates in a posttranslationally translocated competent state (see above), Kar2p in the ER lumen, together with two of the three J-domain-containing proteins, Scj1p and Jem1p, maintain the solubility of ERAD substrates for retrotranslocation; if these chaperones are disabled, substrates aggregate in the ER [55, 180]. It is noteworthy that Sec63p plays only a minor role in this process [181, 180], indicating that BiP functions with its ER membrane partner to facilitate translocation, but interacts instead with the alternate cochaperones during retrotranslocation. Most likely, BiP also interacts with other ER lumenal chaperones during the selection of an aberrant polypeptide in the ER: yeast osteosarcoma-9 protein (Yos9p) [182–185], PDI [186] and the ER degradation enhancing α-mannosidase like protein Htm1p/Mnl1p (EDEM in mammals) [187–189] help select misfolded lumenal proteins for ERAD, and act either upstream or downstream of BiP during the selection process. How these components cooperate with one another to recognize and deliver proteins across the ER membrane to the cytosol remains to be worked out.

BiP is also critical for ERAD in mammals, although in the absence of ERAD-specific mutants in the gene-encoding mammalian BiP, it has been hard to demonstrate this fact unequivocally. But, the regions within a misfolded domain in an ERAD substrate that are recognized by BiP and the regions that trigger its degradation were shown to overlap [190]. In addition, the half-life of an ERAD substrate correlates with its release from BiP [191], and the dissociation of an ERAD substrate from BiP and its subsequent degradation were found to be coincident [192]. It has also been found that a nascent protein that is unable to fold may be passed from calnexin/calreticulin back to BiP prior to its degradation [193].

## V. Unanswered Questions

Although the components and even one structure of the translocation machinery are now available, many mechanistic questions remain. For example, it is not obvious how BiP/Kar2p facilitates translocation. During posttranslational translocation, does this chaperone act as a motor, a ratchet, a combination of the two, and/or does it directly gate the translocon? During cotranslational translocation, does BiP/Kar2p generate a driving force? And, is BiP-mediated sealing of the translocon essential? What signals the release of BiP from the translocon? Does the elongating nascent chain push BiP out of the way, or are one of the many J-domain-containing proteins in the mammalian ER required for gating and for BiP's interaction with the translocon? How is the permeability barrier

maintained during posttranslational translocation? And finally, does BiP/Kar2p regulate channel dynamics in yeast?

Finally, it is important to emphasize that translocation is not an independent event but is only one of many sequential and tightly coupled events during protein biogenesis. Specifically, protein folding, chaperone interactions, and posttranslational modifications all begin during translocation. Therefore, it is important to ask if and how defects in any one of these processes affect translocation efficiency. And in a related vein, how are substrates transferred between the many protein-containing machines in the ER? This question is especially pertinent with regard to ERAD, as this field is in its relative infancy. Not only conventional biochemical and genetic studies but also biophysical and mathematical analyses will be needed to fully understand the many functions of BiP/Kar2p in the ER.

## Acknowledgments

Work in the Brodsky lab was funded by grants from the National Institute of Health, the National Science Foundation, and the Cystic Fibrosis Foundation. K.N. acknowledges the support of the Uehara Memorial Foundation and the American Heart Association.

## References

1. Macario, A.J., Lange, M., Ahring, B.K., and De Macario, E.C. (1999). Stress genes and proteins in the archaea. *Microbiol Mol Biol Rev* 63:923–967.
2. Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci* 62:670–684.
3. Flynn, G.C., Pohl, J., Flocco, M.T., and Rothman, J.E. (1991). Peptide-binding specificity of the molecular chaperone BiP. *Nature* 353:726–730.
4. Blond-Elguindi, S., Cwirla, S.E., Dower, W.J., Lipshutz, R.J., Sprang, S.R., Sambrook, J. F., and Gething, M.J. (1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell* 75:717–728.
5. Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J* 16:1501–1507.
6. Werner-Washburne, M., and Craig, E.A. (1989). Expression of members of the *Saccharomyces cerevisiae* hsp70 multigene family. *Genome* 31:684–689.
7. Cyr, D.M., Langer, T., and Douglas, M.G. (1994). DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. *Trends Biochem Sci* 19:176–181.
8. Cheetham, M.E., and Caplan, A.J. (1998). Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function. *Cell Stress Chaperones* 3:28–36.
9. Craig, E.A., Huang, P., Aron, R., and Andrew, A. (2006). The diverse roles of J-proteins, the obligate Hsp70 co-chaperone. *Rev Physiol Biochem Pharmacol* 156:1–21.

10. Flaherty, K.M., DeLuca-Flaherty, C., and McKay, D.B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346:623–628.
11. Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E., and Hendrickson, W.A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* 272:1606–1614.
12. Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21:932–939.
13. Mayer, M.P., Rudiger, S., and Bukau, B. (2000). Molecular basis for interactions of the DnaK chaperone with substrates. *Biol Chem* 381:877–885.
14. Misselwitz, B., Staeck, O., and Rapoport, T.A. (1998). J proteins catalytically activate Hsp70 molecules to trap a wide range of peptide sequences. *Mol Cell* 2:593–603.
15. Harrison, C.J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., and Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* 276:431–435.
16. Shomura, Y., Dragovic, Z., Chang, H.C., Tzvetkov, N., Young, J.C., Brodsky, J.L., Guerriero, V., Hartl, F.U., and Bracher, A. (2005). Regulation of Hsp70 function by HspBP1: structural analysis reveals an alternate mechanism for Hsp70 nucleotide exchange. *Mol Cell* 17:367–379.
17. Jiang, J., Prasad, K., Lafer, E.M., and Sousa, R. (2005). Structural basis of interdomain communication in the Hsc70 chaperone. *Mol Cell* 20:513–524.
18. Swain, J.F., Schulz, E.G., and Gierasch, L.M. (2006). Direct comparison of a stable isolated Hsp70 substrate-binding domain in the empty and substrate-bound states. *J Biol Chem* 281:1605–1611.
19. Revington, M., Zhang, Y., Yip, G.N., Kurochkin, A.V., and Zuiderweg, E.R. (2005). NMR investigations of allosteric processes in a two-domain *Thermus thermophilus* Hsp70 molecular chaperone. *J Mol Biol* 349:163–183.
19a. Swain, J.F., Dinler, G., Sivendran, R., Montogomery, D.L., Stotz, M., and Gierasch, L. M. (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol Cell* 26:27–39.
20. Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B., and Hartl, F.U. (1994). The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system DnaK, DnaJ, and GrpE. *Proc Natl Acad Sci USA* 91:10345–10349.
21. Buchberger, A., Schroder, H., Hesterkamp, T., Schonfeld, H.J., and Bukau, B. (1996). Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding. *J Mol Biol* 261:328–333.
22. Prasad, K., Heuser, J., Eisenberg, E., and Greene, L. (1994). Complex formation between clathrin and uncoating ATPase. *J Biol Chem* 269:6931–6939.
23. Schmid, D., Baici, A., Gehring, H., and Christen, P. (1994). Kinetics of molecular chaperone action. *Science* 263:971–973.
24. Palleros, D.R., Shi, L., Reid, K.L., and Fink, A.L. (1994). hsp70-protein complexes. Complex stability and conformation of bound substrate protein. *J Biol Chem* 269:13107–13114.
25. McCarty, J.S., Buchberger, A., Reinstein, J., and Bukau, B. (1995). The role of ATP in the functional cycle of the DnaK chaperone system. *J Mol Biol* 249:126–137.
26. Pellecchia, M., Montgomery, D.L., Stevens, S.Y., Vander Kooi, C.W., Feng, H.P., Gierasch, L.M., and Zuiderweg, E.R. (2000). Structural insights into substrate binding by the molecular chaperone DnaK. *Nat Struct Biol* 7:298–303.

27. Russell, R., Wali Karzai, A., Mehl, A.F., and McMacken, R. (1999). DnaJ dramatically stimulates ATP hydrolysis by DnaK: insight into targeting of Hsp70 proteins to polypeptide substrates. *Biochemistry* 38:4165–4176.
28. Gassler, C.S., Buchberger, A., Laufen, T., Mayer, M.P., Schroder, H., Valencia, A., and Bukau, B. (1998). Mutations in the DnaK chaperone affecting interaction with the DnaJ cochaperone. *Proc Natl Acad Sci USA* 95:15229–15234.
29. Suh, W.C., Burkholder, W.F., Lu, C.Z., Zhao, X., Gottesman, M.E., and Gross, C.A. (1998). Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proc Natl Acad Sci USA* 95:15223–15228.
30. Rudiger, S., Schneider-Mergener, J., and Bukau, B. (2001). Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J* 20:1042–1050.
31. Wickner, S., Hoskins, J., and McKenney, K. (1991). Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA. *Nature* 350:165–167.
32. Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., and Hartl, F.U. (1992). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* 356:683–689.
33. Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J., and Bukau, B. (1999). Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. *Proc Natl Acad Sci USA* 96:5452–5457.
34. Han, W., and Christen, P. (2003). Mechanism of the targeting action of DnaJ in the DnaK molecular chaperone system. *J Biol Chem* 278:19038–19043.
35. Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J.S., Rudiger, S., Schonfeld, H.J., Schirra, C., Bujard, H., and Bukau, B. (1996). A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor sigma32. *EMBO J* 15:607–617.
36. Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C., and Zylicz, M. (1991). *Escherichia coli* DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc Natl Acad Sci USA* 88:2874–2878.
37. Haas, I.G., and Wabl, M. (1983). Immunoglobulin heavy chain binding protein. *Nature* 306:387–389.
38. Pouyssegur, J., Shiu, R.P., and Pastan, I. (1977). Induction of two transformation-sensitive membrane polypeptides in normal fibroblasts by a block in glycoprotein synthesis or glucose deprivation. *Cell* 11:941–947.
39. Gething, M.J. (1999). Role and regulation of the ER chaperone BiP. *Semin Cell Dev Biol* 10:465–472.
40. Munro, S., and Pelham, H.R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48:899–907.
41. Semenza, J.C., Hardwick, K.G., Dean, N., and Pelham, H.R. (1990). ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* 61:1349–1357.
42. Rose, M.D., Misra, L.M., and Vogel, J.P. (1989). KAR2, a karyogamy gene, is the yeast homolog of the mammalian BiP/GRP78 gene. *Cell* 57:1211–1221.
43. Ng, D.T., and Walter, P. (1996). ER membrane protein complex required for nuclear fusion. *J Cell Biol* 132:499–509.
44. Latterich, M., and Schekman, R. (1994). The karyogamy gene KAR2 and novel proteins are required for ER-membrane fusion. *Cell* 78:87–98.
45. Normington, K., Kohno, K., Kozutsumi, Y., Gething, M.J., and Sambrook, J. (1989). *S. cerevisiae* encodes an essential protein homologous in sequence and function to mammalian BiP. *Cell* 57:1223–1236.

46. Brizzio, V., Khalfan, W., Huddler, D., Beh, C.T., Andersen, S.S., Latterich, M., and Rose, M.D. (1999). Genetic interactions between KAR7/SEC71, KAR8/JEM1, KAR5, and KAR2 during nuclear fusion in *Saccharomyces cerevisiae*. *Mol Biol Cell* 10:609–626.
47. Feldheim, D., Rothblatt, J., and Schekman, R. (1992). Topology and functional domains of Sec63p, an endoplasmic reticulum membrane protein required for secretory protein translocation. *Mol Cell Biol* 12:3288–3296.
48. Schlenstedt, G., Harris, S., Risse, B., Lill, R., and Silver, P.A. (1995). A yeast DnaJ homologue, Scj1p, can function in the endoplasmic reticulum with BiP/Kar2p via a conserved domain that specifies interactions with Hsp70s. *J Cell Biol* 129:979–988.
49. Nishikawa, S., and Endo, T. (1997). The yeast JEM1p is a DnaJ-like protein of the endoplasmic reticulum membrane required for nuclear fusion. *J Biol Chem* 272:12889–12892.
50. Nishikawa, S., and Endo, T. (1998). Reinvestigation of the functions of the hydrophobic segment of Jem1p, a yeast endoplasmic reticulum membrane protein mediating nuclear fusion. *Biochem Biophys Res Commun* 244:785–789.
51. Matlack, K.E., Misselwitz, B., Plath, K., and Rapoport, T.A. (1999). BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* 97:553–564.
52. Misselwitz, B., Staeck, O., Matlack, K.E., and Rapoport, T.A. (1999). Interaction of BiP with the J-domain of the Sec63p component of the endoplasmic reticulum protein translocation complex. *J Biol Chem* 274:20110–20115.
53. Brodsky, J.L., and Schekman, R. (1993). A Sec63p-BiP complex from yeast is required for protein translocation in a reconstituted proteoliposome. *J Cell Biol* 123:1355–1363.
54. Corsi, A.K., and Schekman, R. (1997). The lumenal domain of Sec63p stimulates the ATPase activity of BiP and mediates BiP recruitment to the translocon in *Saccharomyces cerevisiae*. *J Cell Biol* 137:1483–1493.
55. Nishikawa, S.I., Fewell, S.W., Kato, Y., Brodsky, J.L., and Endo, T. (2001). Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. *J Cell Biol* 153:1061–1070.
56. Brightman, S.E., Blatch, G.L., and Zetter, B.R. (1995). Isolation of a mouse cDNA encoding MTJ1, a new murine member of the DnaJ family of proteins. *Gene* 153:249–254.
57. Dudek, J., Volkmer, J., Bies, C., Guth, S., Muller, A., Lerner, M., Feick, P., Schafer, K.H., Morgenstern, E., Hennessy, F., Blatch, G.L., Janoscheck, K., *et al.* (2002). A novel type of co-chaperone mediates transmembrane recruitment of DnaK-like chaperones to ribosomes. *EMBO J* 21:2958–2967.
58. Chevalier, M., Rhee, H., Elguindi, E.C., and Blond, S.Y. (2000). Interaction of murine BiP/GRP78 with the DnaJ homologue MTJ1. *J Biol Chem* 275:19620–19627.
59. Kroczynska, B., Evangelista, C.M., Samant, S.S., Elguindi, E.C., and Blond, S.Y. (2004). The SANT2 domain of the murine tumor cell DnaJ-like protein 1 human homologue interacts with alpha1-antichymotrypsin and kinetically interferes with its serpin inhibitory activity. *J Biol Chem* 279:11432–11443.
60. Zupicich, J., Brenner, S.E., and Skarnes, W.C. (2001). Computational prediction of membrane-tethered transcription factors. *Genome Biol* 2:RESEARCH0050.
61. Skowronek, M.H., Rotter, M., and Haas, I.G. (1999). Molecular characterization of a novel mammalian DnaJ-like Sec63p homolog. *Biol Chem* 380:1133–1138.
62. Tyedmers, J., Lerner, M., Bies, C., Dudek, J., Skowronek, M.H., Haas, I.G., Heim, N., Nastainczyk, W., Volkmer, J., and Zimmermann, R. (2000). Homologs of the yeast Sec complex subunits Sec62p and Sec63p are abundant proteins in dog pancreas microsomes. *Proc Natl Acad Sci USA* 97:7214–7219.

63. Bies, C., Guth, S., Janoschek, K., Nastainczyk, W., Volkmer, J., and Zimmermann, R. (1999). A Scj1p homolog and folding catalysts present in dog pancreas microsomes. *Biol Chem* 380:1175–1182.
64. Yu, M., Haslam, R.H., and Haslam, D.B. (2000). HEDJ, an Hsp40 co-chaperone localized to the endoplasmic reticulum of human cells. *J Biol Chem* 275:24984–24992.
65. Lau, P.P., Villanueva, H., Kobayashi, K., Nakamuta, M., Chang, B.H., and Chan, L. (2001). A DnaJ protein, apobec-1-binding protein-2, modulates apolipoprotein B mRNA editing. *J Biol Chem* 276:46445–46452.
66. Prols, F., Mayer, M.P., Renner, O., Czarnecki, P.G., Ast, M., Gassler, C., Wilting, J., Kurz, H., and Christ, B. (2001). Upregulation of the cochaperone Mdg1 in endothelial cells is induced by stress and during *in vitro* angiogenesis. *Exp Cell Res* 269:42–53.
67. Shen, Y., Meunier, L., and Hendershot, L.M. (2002). Identification and characterization of a novel endoplasmic reticulum (ER) DnaJ homologue, which stimulates ATPase activity of BiP *in vitro* and is induced by ER stress. *J Biol Chem* 277:15947–15956.
68. Cunnea, P.M., Miranda-Vizuete, A., Bertoli, G., Simmen, T., Damdimopoulos, A.E., Hermann, S., Leinonen, S., Huikko, M.P., Gustafsson, J.A., Sitia, R., and Spyrou, G. (2003). ERdj5, an endoplasmic reticulum (ER)-resident protein containing DnaJ and thioredoxin domains, is expressed in secretory cells or following ER stress. *J Biol Chem* 278:1059–1066.
69. Hosoda, A., Kimata, Y., Tsuru, A., and Kohno, K. (2003). JPDI, a novel endoplasmic reticulum-resident protein containing both a BiP-interacting J-domain and thioredoxin-like motifs. *J Biol Chem* 278:2669–2676.
70. Blau, M., Mullapudi, S., Becker, T., Dudek, J., Zimmermann, R., Penczek, P.A., and Beckmann, R. (2005). ERj1p uses a universal ribosomal adaptor site to coordinate the 80S ribosome at the membrane. *Nat Struct Mol Biol* 12:1015–1016.
71. Dudek, J., Greiner, M., Muller, A., Hendershot, L.M., Kopsch, K., Nastainczyk, W., and Zimmermann, R. (2005). ERj1p has a basic role in protein biogenesis at the endoplasmic reticulum. *Nat Struct Mol Biol* 12:1008–1014.
72. Shen, Y., and Hendershot, L.M. (2005). ERdj3, a stress-inducible endoplasmic reticulum DnaJ homologue, serves as a cofactor for BiP's interactions with unfolded substrates. *Mol Biol Cell* 16:40–50.
72a. Brodsky, J.L., and Bracher, A. (2007). Nucleotide exchange factors for Hsp70 molecular chaperones, in G.L. Blatch (ed.), *Networking of Chaperones by Co-Chaperones*, pp. 1–12, www.Eurekah.com, Georgetown, Tx: Landes Biosciences.
73. Chung, K.T., Shen, Y., and Hendershot, L.M. (2002). BAP, a mammalian BiP-associated protein, is a nucleotide exchange factor that regulates the ATPase activity of BiP. *J Biol Chem* 277:47557–47563.
74. Kabani, M., Beckerich, J.M., and Gaillardin, C. (2000). Sls1p stimulates Sec63p-mediated activation of Kar2p in a conformation-dependent manner in the yeast endoplasmic reticulum. *Mol Cell Biol* 20:6923–6934.
75. Chen, X., Easton, D., Oh, H.J., Lee-Yoon, D.S., Liu, X., and Subjeck, J. (1996). The 170 kDa glucose regulated stress protein is a large HSP70-, HSP110-like protein of the endoplasmic reticulum. *FEBS Lett* 380:68–72.
76. Steel, G.J., Fullerton, D.M., Tyson, J.R., and Stirling, C.J. (2004). Coordinated activation of Hsp70 chaperones. *Science* 303:98–101.
77. Blond-Elguindi, S., Fourie, A.M., Sambrook, J.F., and Gething, M.J. (1993). Peptide-dependent stimulation of the ATPase activity of the molecular chaperone BiP is the result of conversion of oligomers to active monomers. *J Biol Chem* 268:12730–12735.
78. Freiden, P.J., Gaut, J.R., and Hendershot, L.M. (1992). Interconversion of three differentially modified and assembled forms of BiP. *EMBO J* 11:63–70.

79. Lievremont, J.P., Rizzuto, R., Hendershot, L., and Meldolesi, J. (1997). BiP, a major chaperone protein of the endoplasmic reticulum lumen, plays a direct and important role in the storage of the rapidly exchanging pool of $Ca2^{+}$. *J Biol Chem* 272:30873–30879.
80. Lamb, H.K., Mee, C., Xu, W., Liu, L., Blond, S., Cooper, A., Charles, I.G., and Hawkins, A.R. (2006). The affinity of a major $Ca2^{+}$ binding site on GRP78 is differentially enhanced by ADP and ATP. *J Biol Chem* 281:8796–8805.
81. Mori, K., Sant, A., Kohno, K., Normington, K., Gething, M.J., and Sambrook, J.F. (1992). A 22 bp cis-acting element is necessary and sufficient for the induction of the yeast KAR2 (BiP) gene by unfolded proteins. *EMBO J* 11:2583–2593.
82. Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., and Walter, P. (2000). Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* 101:249–258.
83. Schroder, M., and Kaufman, R.J. (2005). The mammalian unfolded protein response. *Annu Rev Biochem* 74:739–789.
84. Rapoport, T.A., Matlack, K.E., Plath, K., Misselwitz, B., and Staeck, O. (1999). Posttranslational protein translocation across the membrane of the endoplasmic reticulum. *Biol Chem* 380:1143–1150.
85. Brodsky, J.L. (1996). Post-translational protein translocation: not all hsc70s are created equal. *Trends Biochem Sci* 21:122–126.
86. Walter, P., and Johnson, A.E. (1994). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu Rev Cell Biol* 10:87–119.
87. Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* 134:269–278.
88. Matlack, K.E., Mothes, W., and Rapoport, T.A. (1998). Protein translocation: tunnel vision. *Cell* 92:381–390.
89. Johnson, A.E., and van Waes, M.A. (1999). The translocon: a dynamic gateway at the ER membrane. *Annu Rev Cell Dev Biol* 15:799–842.
90. Simon, S.M., and Blobel, G. (1991). A protein-conducting channel in the endoplasmic reticulum. *Cell* 65:371–380.
91. Gorlich, D., and Rapoport, T.A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75:615–630.
92. Hanein, D., Matlack, K.E., Jungnickel, B., Plath, K., Kalies, K.U., Miller, K.R., Rapoport, T.A., and Akey, C.W. (1996). Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* 87:721–732.
93. Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997). Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* 278:2123–2126.
94. Deshaies, R.J., Sanders, S.L., Feldheim, D.A., and Schekman, R. (1991). Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. *Nature* 349:806–808.
95. Wilkinson, B.M., Critchley, A.J., and Stirling, C.J. (1996). Determination of the transmembrane topology of yeast Sec61p, an essential component of the endoplasmic reticulum translocation complex. *J Biol Chem* 271:25590–25597.
96. Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *EMBO J* 15:1482–1494.
97. Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T.A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81:561–570.

98. Wittke, S., Dunnwald, M., and Johnsson, N. (2000). Sec62p, a component of the endoplasmic reticulum protein translocation machinery, contains multiple binding sites for the Sec-complex. *Mol Biol Cell* 11:3859–3871.
99. Willer, M., Jermy, A.J., Young, B.P., and Stirling, C.J. (2003). Identification of novel protein-protein interactions at the cytosolic surface of the Sec63 complex in the yeast ER membrane. *Yeast* 20:133–148.
100. Ponting, C.P. (2000). Proteins of the endoplasmic-reticulum-associated degradation pathway: domain detection and function prediction. *Biochem J* 351(Pt. 2):527–535.
101. Noble, S.M., and Guthrie, C. (1996). Transcriptional pulse-chase analysis reveals a role for a novel snRNP-associated protein in the manufacture of spliceosomal snRNPs. *EMBO J* 15:4368–4379.
102. Meyer, H.A., Grau, H., Kraft, R., Kostka, S., Prehn, S., Kalies, K.U., and Hartmann, E. (2000). Mammalian Sec61 is associated with Sec62 and Sec63. *J Biol Chem* 275:14550–14557.
103. Jermy, A.J., Willer, M., Davis, E., Wilkinson, B.M., and Stirling, C.J. (2006). The Brl domain in Sec63p is required for assembly of functional endoplasmic reticulum translocons. *J Biol Chem* 281:7899–7906.
104. Deshaies, R.J., and Schekman, R. (1989). SEC62 encodes a putative membrane protein required for protein translocation into the yeast endoplasmic reticulum. *J Cell Biol* 109:2653–2664.
105. Menetret, J.F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T.A., and Akey, C.W. (2000). The structure of ribosome-channel complexes engaged in protein translocation. *Mol Cell* 6:1219–1232.
106. Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. *Cell* 107:361–372.
107. Manting, E.H., van Der Does, C., Remigy, H., Engel, A., and Driessen, A.J. (2000). SecYEG assembles into a tetramer to form the active protein translocation channel. *EMBO J* 19:852–861.
108. Bessonneau, P., Besson, V., Collinson, I., and Duong, F. (2002). The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *EMBO J* 21:995–1003.
109. Hamman, B.D., Chen, J.C., Johnson, E.E., and Johnson, A.E. (1997). The aqueous pore through the translocon has a diameter of 40–60 A during cotranslational protein translocation at the ER membrane. *Cell* 89:535–544.
110. Van den Berg, B., Clemons, W.M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427:36–44.
111. Zimmermann, R., Sagstetter, M., Lewis, M.J., and Pelham, H.R. (1988). Seventy-kilodalton heat shock proteins and an additional component from reticulocyte lysate stimulate import of M13 procoat protein into microsomes. *EMBO J* 7:2875–2880.
112. Hann, B.C., and Walter, P. (1991). The signal recognition particle in *S. cerevisiae*. *Cell* 67:131–144.
113. Brown, J.D., Hann, B.C., Medzihradszky, K.F., Niwa, M., Burlingame, A.L., and Walter, P. (1994). Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. *EMBO J* 13:4390–4400.
114. Wilkinson, B.M., Regnacq, M., and Stirling, C.J. (1997). Protein translocation across the membrane of the endoplasmic reticulum. *J Membr Biol* 155:189–197.
115. Ngosuwan, J., Wang, N.M., Fung, K.L., and Chirico, W.J. (2003). Roles of cytosolic Hsp70 and Hsp40 molecular chaperones in post-translational translocation of presecretory proteins into the endoplasmic reticulum. *J Biol Chem* 278:7034–7042.

116. Plath, K., and Rapoport, T.A. (2000). Spontaneous release of cytosolic proteins from posttranslational substrates before their transport into the endoplasmic reticulum. *J Cell Biol* 151:167–178.
117. Matlack, K.E., Plath, K., Misselwitz, B., and Rapoport, T.A. (1997). Protein transport by purified yeast Sec complex and Kar2p without membranes. *Science* 277:938–941.
118. Lyman, S.K., and Schekman, R. (1997). Binding of secretory precursor polypeptides to a translocon subcomplex is regulated by BiP. *Cell* 88:85–96.
119. Plath, K., Mothes, W., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* 94:795–807.
120. Plath, K., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (2004). Interactions between Sec complex and prepro-alpha-factor during posttranslational protein transport into the endoplasmic reticulum. *Mol Biol Cell* 15:1–10.
121. Pilon, M., Romisch, K., Quach, D., and Schekman, R. (1998). Sec61p serves multiple roles in secretory precursor binding and translocation into the endoplasmic reticulum membrane. *Mol Biol Cell* 9:3455–3473.
122. Sanders, S.L., Whitfield, K.M., Vogel, J.P., Rose, M.D., and Schekman, R.W. (1992). Sec61p and BiP directly facilitate polypeptide translocation into the ER. *Cell* 69:353–365.
123. Simon, S.M., Peskin, C.S., and Oster, G.F. (1992). What drives the translocation of proteins? *Proc Natl Acad Sci USA* 89:3770–3774.
124. Schneider, H.C., Berthold, J., Bauer, M.F., Dietmeier, K., Guiard, B., Brunner, M., and Neupert, W. (1994). Mitochondrial Hsp70/MIM44 complex facilitates protein import. *Nature* 371:768–774.
125. Liebermeister, W., Rapoport, T.A., and Heinrich, R. (2001). Ratcheting in post-translational protein translocation: A mathematical model. *J Mol Biol* 305:643–656.
126. De Los Rios, P., Ben-Zvi, A., Slutsky, O., Azem, A., and Goloubinoff, P. (2006). Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc Natl Acad Sci USA* 103:6166–6171.
127. Glick, B.S., Wachter, C., Reid, G.A., and Schatz, G. (1993). Import of cytochrome b2 to the mitochondrial intermembrane space: the tightly folded heme-binding domain makes import dependent upon matrix ATP. *Protein Sci* 2:1901–1917.
128. Glick, B.S. (1995). Can Hsp70 proteins act as force-generating motors? *Cell* 80:11–14.
129. Hannavy, K., Rospert, S., and Schatz, G. (1993). Protein import into mitochondria: a paradigm for the translocation of polypeptides across membranes. *Curr Opin Cell Biol* 5:694–700.
130. Jungnickel, B., Rapoport, T.A., and Hartmann, E. (1994). Protein translocation: common themes from bacteria to man. *FEBS Lett* 346:73–77.
131. Voisine, C., Craig, E.A., Zufall, N., von Ahsen, O., Pfanner, N., and Voos, W. (1999). The protein import motor of mitochondria: unfolding and trapping of preproteins are distinct and separable functions of matrix Hsp70. *Cell* 97:565–574.
132. Hwang, S.T., Wachter, C., and Schatz, G. (1991). Protein import into the yeast mitochondrial matrix. A new translocation intermediate between the two mitochondrial membranes. *J Biol Chem* 266:21083–21089.
133. Chauwin, J.F., Oster, G., and Glick, B.S. (1998). Strong precursor-pore interactions constrain models for mitochondrial protein import. *Biophys J* 74:1732–1743.
134. McClellan, A.J., Endres, J.B., Vogel, J.P., Palazzi, D., Rose, M.D., and Brodsky, J.L. (1998). Specific molecular chaperone interactions and an ATP-dependent conformational change are required during posttranslational protein translocation into the yeast ER. *Mol Biol Cell* 9:3533–3545.

135. Tyson, J.R., and Stirling, C.J. (2000). LHS1 and SIL1 provide a lumenal function that is essential for protein translocation into the endoplasmic reticulum. *EMBO J* 19:6440–6452.
136. Hamilton, T.G., and Flynn, G.C. (1996). Cer1p, a novel Hsp70-related protein required for posttranslational endoplasmic reticulum translocation in yeast. *J Biol Chem* 271:30610–30613.
137. Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. *Annu Rev Biochem* 70:755–775.
138. Miller, J.D., Wilhelm, H., Gierasch, L., Gilmore, R., and Walter, P. (1993). GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature* 366:351–354.
139. Powers, T., and Walter, P. (1995). Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science* 269:1422–1424.
140. Bacher, G., Lutcke, H., Jungnickel, B., Rapoport, T.A., and Dobberstein, B. (1996). Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature* 381:248–251.
141. Millman, J.S., and Andrews, D.W. (1997). Switching the model: a concerted mechanism for GTPases in protein targeting. *Cell* 89:673–676.
142. Rapiejko, P.J., and Gilmore, R. (1997). Empty site forms of the SRP54 and SR alpha GTPases mediate targeting of ribosome-nascent chain complexes to the endoplasmic reticulum. *Cell* 89:703–713.
143. Song, W., Raden, D., Mandon, E.C., and Gilmore, R. (2000). Role of Sec61alpha in the regulated transfer of the ribosome-nascent chain complex from the signal recognition particle to the translocation channel. *Cell* 100:333–343.
144. Kida, Y., Mihara, K., and Sakaguchi, M. (2005). Translocation of a long amino-terminal domain through ER membrane by following signal-anchor sequence. *EMBO J* 24:3202–3213.
145. Nicchitta, C.V., and Blobel, G. (1993). Lumenal proteins of the mammalian endoplasmic reticulum are required to complete protein translocation. *Cell* 73:989–998.
146. Klappa, P., Mayinger, P., Pipkorn, R., Zimmermann, M., and Zimmermann, R. (1991). A microsomal protein is involved in ATP-dependent transport of presecretory proteins into mammalian microsomes. *EMBO J* 10:2795–2803.
147. Liao, S., Lin, J., Do, H., and Johnson, A.E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell* 90:31–41.
148. Dierks, T., Volkmer, J., Schlenstedt, G., Jung, C., Sandholzer, U., Zachmann, K., Schlotterhose, P., Neifer, K., Schmidt, B., and Zimmermann, R. (1996). A microsomal ATP-binding protein involved in efficient protein transport into the mammalian endoplasmic reticulum. *EMBO J* 15:6931–6942.
149. Tyedmers, J., Lerner, M., Wiedmann, M., Volkmer, J., and Zimmermann, R. (2003). Polypeptide-binding proteins mediate completion of co-translational protein translocation into the mammalian endoplasmic reticulum. *EMBO Rep* 4:505–510.
150. Hamman, B.D., Hendershot, L.M., and Johnson, A.E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the lumenal end of the translocon pore before and early in translocation. *Cell* 92:747–758.
151. Young, B.P., Craven, R.A., Reid, P.J., Willer, M., and Stirling, C.J. (2001). Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum *in vivo*. *EMBO J* 20:262–271.

152. Willer, M., Jermy, A.J., Steel, G.J., Garside, H.J., Carter, S., and Stirling, C.J. (2003). An *in vitro* assay using overexpressed yeast SRP demonstrates that cotranslational translocation is dependent upon the J-domain of Sec63p. *Biochemistry* 42:7171–7177.
153. Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc Natl Acad Sci USA* 92:9643–9646.
154. Wirth, A., Jung, M., Bies, C., Frien, M., Tyedmers, J., Zimmermann, R., and Wagner, R. (2003). The Sec61p complex is a dynamic precursor activated channel. *Mol Cell* 12:261–268.
155. Crowley, K.S., Liao, S., Worrell, V.E., Reinhart, G.D., and Johnson, A.E. (1994). Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* 78:461–471.
156. Alder, N.N., Shen, Y., Brodsky, J.L., Hendershot, L.M., and Johnson, A.E. (2005). The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *J Cell Biol* 168:389–399.
157. Rapoport, T.A., Goder, V., Heinrich, S.U., and Matlack, K.E. (2004). Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol* 14:568–575.
158. Gumbart, J., and Schulten, K. (2006). Molecular dynamics studies of the archaeal translocon. *Biophys J* 90:2356–2367.
159. Netzer, W.J., and Hartl, F.U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388:343–349.
160. Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol Cell* 10:769–778.
161. Kleizen, B., van Vlijmen, T., de Jonge, H.R., and Braakman, I. (2005). Folding of CFTR is predominantly cotranslational. *Mol Cell* 20:277–287.
162. Du, K., Sharma, M., and Lukacs, G.L. (2005). The DeltaF508 cystic fibrosis mutation impairs domain-domain interactions and arrests post-translational folding of CFTR. *Nat Struct Mol Biol* 12:17–25.
163. Meunier, L., Usherwood, Y.K., Chung, K.T., and Hendershot, L.M. (2002). A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol Biol Cell* 13:4456–4469.
164. Simons, J.F., Ferro-Novick, S., Rose, M.D., and Helenius, A. (1995). BiP/Kar2p serves as a molecular chaperone during carboxypeptidase Y folding in yeast. *J Cell Biol* 130:41–49.
165. Williams, D.B. (2006). Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J Cell Sci* 119:615–623.
166. Ellgaard, L., and Helenius, A. (2001). ER quality control: towards an understanding at the molecular level. *Curr Opin Cell Biol* 13:431–437.
167. Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* 4:181–191.
168. Spiro, R.G. (2004). Role of N-linked polymannose oligosaccharides in targeting glycoproteins for endoplasmic reticulum-associated degradation. *Cell Mol Life Sci* 61:1025–1041.
169. Helenius, A., and Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 73:1019–1049.
170. Parodi, A.J. (2000). Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *Biochem J* 348:1–13.
171. Molinari, M., and Helenius, A. (2000). Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* 288:331–333.

172. Wang, N., Daniels, R., and Hebert, D.N. (2005). The cotranslational maturation of the type I membrane glycoprotein tyrosinase: the heat shock protein 70 system hands off to the lectin-based chaperone system. *Mol Biol Cell* 16:3740–3752.
173. Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005). ERAD: the long road to destruction. *Nat Cell Biol* 7:766–772.
174. Hampton, R.Y. (2002). ER-associated degradation in protein quality control and cellular regulation. *Curr Opin Cell Biol* 14:476–482.
175. Kostova, Z., and Wolf, D.H. (2003). For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *EMBO J* 22:2309–2317.
176. McCracken, A.A., and Brodsky, J.L. (2003). Evolving questions and paradigm shifts in endoplasmic-reticulum-associated degradation (ERAD). *Bioessays* 25:868–877.
177. Sayeed, A., and Ng, D.T. (2005). Search and destroy: ER quality control and ER-associated protein degradation. *Crit Rev Biochem Mol Biol* 40:75–91.
178. Ye, Y., Shibata, Y., Yun, C., Ron, D., and Rapoport, T.A. (2004). A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* 429:841–847.
179. Lilley, B.N., and Ploegh, H.L. (2004). A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 429:834–840.
180. Nishikawa, S., Brodsky, J.L., and Nakatsukasa, K. (2005). Roles of molecular chaperones in endoplasmic reticulum (ER) quality control and ER-associated degradation (ERAD). *J Biochem (Tokyo)* 137:551–555.
181. Plemper, R.K., Bohmler, S., Bordallo, J., Sommer, T., and Wolf, D.H. (1997). Mutant analysis links the translocon and BiP to retrograde protein transport for ER degradation. *Nature* 388:891–895.
182. Kim, W., Spear, E.D., and Ng, D.T. (2005). Yos9p detects and targets misfolded glycoproteins for ER-associated degradation. *Mol Cell* 19:753–764.
183. Bhamidipati, A., Denic, V., Quan, E.M., and Weissman, J.S. (2005). Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen. *Mol Cell* 19:741–751.
184. Buschhorn, B.A., Kostova, Z., Medicherla, B., and Wolf, D.H. (2004). A genome-wide screen identifies Yos9p as essential for ER-associated degradation of glycoproteins. *FEBS Lett* 577:422–426.
185. Szathmary, R., Bielmann, R., Nita-Lazar, M., Burda, P., and Jakob, C.A. (2005). Yos9 protein is essential for degradation of misfolded glycoproteins and may function as lectin in ERAD. *Mol Cell* 19:765–775.
186. Gillece, P., Luz, J.M., Lennarz, W.J., de La Cruz, F.J., and Romisch, K. (1999). Export of a cysteine-free misfolded secretory protein from the endoplasmic reticulum for degradation requires interaction with protein disulfide isomerase. *J Cell Biol* 147:1443–1456.
187. Jakob, C.A., Bodmer, D., Spirig, U., Battig, P., Marcil, A., Dignard, D., Bergeron, J.J., Thomas, D.Y., and Aebi, M. (2001). Htm1p, a mannosidase-like protein, is involved in glycoprotein degradation in yeast. *EMBO Rep* 2:423–430.
188. Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L.O., Herscovics, A., and Nagata, K. (2001). A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. *EMBO Rep* 2:415–422.
189. Nakatsukasa, K., Nishikawa, S., Hosokawa, N., Nagata, K., and Endo, T. (2001). Mnl1p, an alpha -mannosidase-like protein in yeast *Saccharomyces cerevisiae*, is required for endoplasmic reticulum-associated degradation of glycoproteins. *J Biol Chem* 276:8635–8638.

190. Schmitz, A., Maintz, M., Kehle, T., and Herzog, V. (1995). *In vivo* iodination of a misfolded proinsulin reveals co-localized signals for Bip binding and for degradation in the ER. *EMBO J* 14:1091–1098.
191. Skowronek, M.H., Hendershot, L.M., and Haas, I.G. (1998). The variable domain of nonassembled Ig light chains determines both their half-life and binding to the chaperone BiP. *Proc Natl Acad Sci USA* 95:1574–1578.
192. Chillaron, J., and Haas, I.G. (2000). Dissociation from BiP and retrotranslocation of unassembled immunoglobulin light chains are tightly coupled to proteasome activity. *Mol Biol Cell* 11:217–226.
193. Molinari, M., Galli, C., Vanoni, O., Arnold, S.M., and Kaufman, R.J. (2005). Persistent glycoprotein misfolding activates the glucosidase II/UGT1-driven calnexin cycle to delay aggregation and loss of folding competence. *Mol Cell* 20:503–512.

This page intentionally left blank

# 11

# *Calnexin, Calreticulin, and Their Associated Oxidoreductase ERp57*

BRADLEY R. PEARSE • DANIEL N. HEBERT

*Department of Biochemistry and Molecular Biology*
*Program in Cellular and Molecular Biology*
*University of Massachusetts*
*Amherst, MA 01003 USA*

## I. Abstract

The lectin chaperones and ERp57 play critical roles in maintaining a high fidelity of glycoprotein folding during the early stages of maturation, as proteins emerge from the translocon channel, as well as later stages of protein quality control. Calnexin and calreticulin associate with monoglucosylated proteins, which support the recruitment of their associated oxidoreductase ERp57 to the maturing protein. Calnexin and calreticulin binding can direct the folding of the nascent chain and protect it from nonproductive aggregate formation. ERp57 catalyzes disulfide bond formation, reduction, and isomerization of nascent substrates. This network of chaperones and the oxidoreductase can also modulate the activity of mature proteins and assist in the presentation of antigen by major histocompatibility (MHC) class I. They are capable of performing these diverse functions because of some fascinating structural characteristics and exquisite active site chemistry. Our understanding of the mechanisms of action of these components has been greatly aided by recently solved crystallographic structures for soluble calnexin and protein disulfide isomerase (PDI). In this chapter, we discuss the mechanisms and roles of calnexin, calreticulin, and their associated oxidoreductase ERp57 in glycoprotein folding and quality control.

THE ENZYMES, Vol. XXV


ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25011-5

## II. Introduction

Proteins destined for the cell surface or the extracellular milieu of eukaryotic cells are synthesized by cytosolic ribosomes that are cotranslationally targeted to the endoplasmic reticulum (ER) membrane. These proteins are cotranslationally translocated across the ER membrane via the Sec61 translocon. The oligosaccharyl transferase (OST) scans the nascent polypeptide for the consensus sequence of Asn-linked glycosylation (Asn-X-Ser/Thr) as the protein emerges from the translocon channel [1, 2]. A 14-member carbohydrate moiety ($Glc_3Man_9GlcNAc_2$, Figure 11.1) is transferred *en bloc* by the OST from a lipid-linked precursor to form an N-glycosidic bond with an Asn side chain of the nascent protein. The majority of the proteins that traffic through the secretory pathway receive multiple Asn-linked glycans [3].

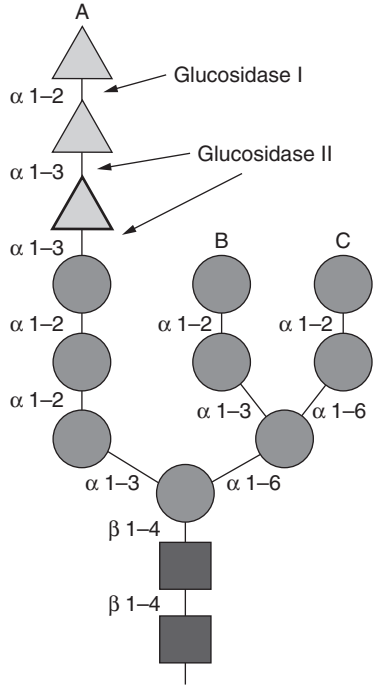



FIG. 11.1. Schematic of an Asn-linked glycan. The Asn-linked glycan is composed of 14 oligosaccharides linked through glycosidic bonds. *N*-acetylglucosamine is depicted as a square, mannose as a circle, and glucose as a triangle. The orientation of the glycosidic bonds are designated. The three mannose branches are labeled A, B, and C, respectively. The glucose residue that is required for lectin chaperone association is bold. The targets of glucosidase I and II are depicted.

The addition of these large hydrophilic modifications dramatically alters the inherent characteristics of the nascent glycoprotein supporting changes in the solubility and hydrodynamic volume of the protein. In addition, these carbohydrates can provide a platform for protein–carbohydrate interactions between the nascent glycoprotein and maturation enzymes that possess carbohydrate-binding activities [1, 2]. A set of lectin chaperones and their associated enzymes help to maximize the efficiency of glycoprotein maturation, as well as monitor the integrity of the maturation process through early cotranslational stages of maturation and latter posttranslational stages of retention, oligomeric assembly, and in the case of substrate malformation, degradation [1, 2]. These functions are performed in part by the ER carbohydrate-binding chaperones calnexin and calreticulin that bind proteins with Asn-linked glycans bearing a single glucose. In this chapter, we will discuss the mechanisms and roles of calnexin, calreticulin, and their associated oxidoreductase ERp57 in glycoprotein folding and quality control.

## III. Structural Characteristics of Calnexin and Calreticulin

Calnexin and calreticulin are lectin chaperones that assist in glycoprotein maturation and quality control in eukaryotic cells. Calnexin is a 592 amino acid type I ER membrane protein with a lumenally exposed N-terminal ectodomain and a short C-terminal cytoplasmic tail [4–6]. Calreticulin is its soluble paralogue of 417 amino acids that possesses 33% sequence identity with calnexin [6–9]. Through analysis of their sequences and structural studies several homologous domains can be identified (Figure 11.2). These domains include several distal regions that fold into a single carbohydrate-binding site (Figure 11.2, gray shaded structures) and the extended Pro-rich repeat regions termed the P-domain (Figure 11.2, numbered open structures). Localization of the carbohydrate-binding sites required solving the crystal structure for the ectodomain of calnexin and mutagenesis studies since no signature carbohydrate-binding domain could be identified from the sequence [10, 11]. The P-domain forms an arm that extends away from the carbohydrate-binding domain and creates a platform for accessory protein binding. These proteins also contain regulatory and ER retention sequences that control the activity and cellular localization of the chaperones.

### A. Carbohydrate-Binding Site

The lectin site was initially proposed to reside in the repeat regions of the extended P-domain in calnexin and calreticulin [12]. However, crystallization and mutagenesis studies indicated that the oligosaccharide-binding site

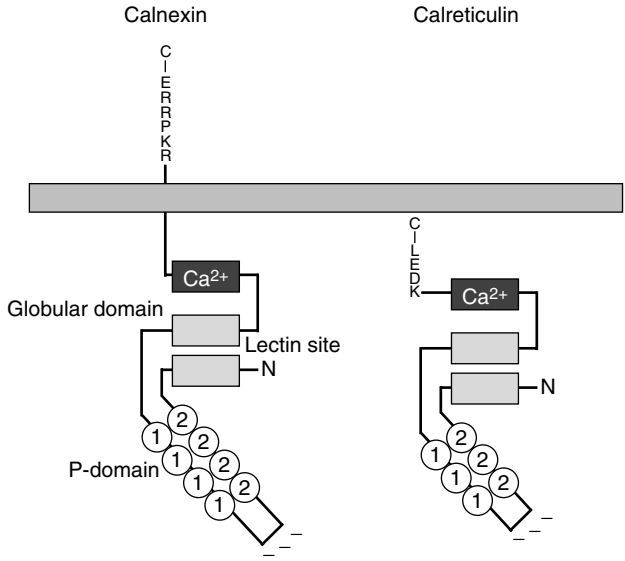



FIG. 11.2. Domain organization of calnexin and calreticulin. Calnexin is a type I membrane-anchored ER protein that contains three distinct regions. The globular domain lectin-binding site, colored gray, results from a fold that joins two distal regions of the polypeptide chain. The $Ca^{2+}$-binding site, colored black, is located near the C-terminus and resides in the globular domain. The numbered repeat motifs are located in the P-domain and reside on either side of the hairpin. The negatively charged tip of the P-domain is indicated and the C-terminal ER retention signal for calnexin (R-K-P-R-R-E) is marked. Calreticulin is the soluble paralogue of calnexin. Its regions are marked in the same manner as that for calnexin. Calreticulin contains only three repeat motifs as opposed to the four for calnexin. The C-terminal retention signal for calreticulin (K-D-E-L) is designated.

resides in the nonrepeat regions [10, 11]. The 2.9-Å crystal structure of canine calnexin revealed two domains, a globular β-stranded region and an extended hairpin domain, which contains the repeating units characteristic of the P-domain. The fold of the globular domain is primarily β-sandwich and contains the majority of the secondary structural elements found within the protein (Figure 11.3) [10]. The structural features of the similar domain in calreticulin have never been visualized; however, mutagenesis studies indicate strong similarities between calnexin and calreticulin lectin domains. The structure of the carbohydrate domain for calnexin is strikingly similar to a number of legume lectins, galectins, and bacterial glycosidases, which satisfies previous concerns about the lack of sequence similarity between the carbohydrate-binding chaperones and the classical lectins [10, 13].

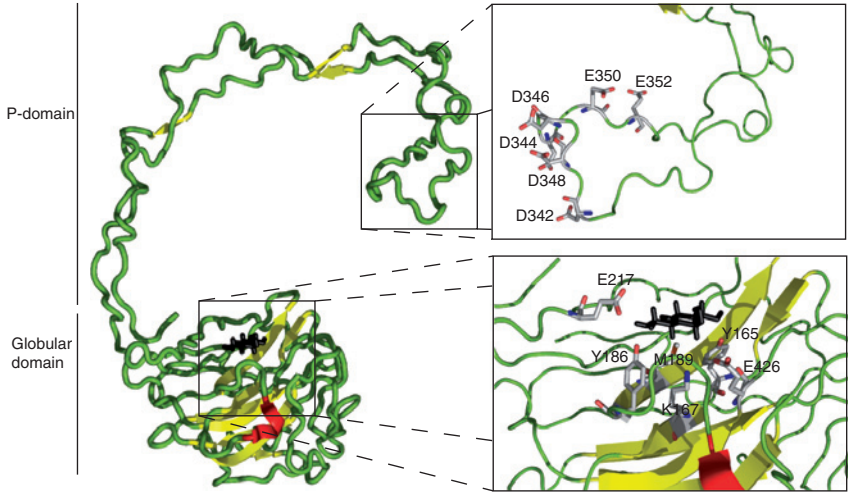



FIG. 11.3. Crystal structure of soluble calnexin and insight into the lectin and ERp57-docking sites. The 2.9-Å crystal structure of soluble canine calnexin (PDB code: 1JHN) is colored by structural elements with β-strands in yellow, α-helices in red, and unstructured loops in green. The general boundaries of the P-domain and the globular domain are indicated with bars. The site of ERp57 docking on the tip of the P-domain is enlarged to display residues that are important for the interaction [19]. Also enlarged is the lectin-binding site with the residues that coordinate the bound glucose highlighted [10]. A hypothetical positioning of glucose is included to emphasize where the oligosaccharide is bound. Labeled residues are colored by elements. Molecular graphics were created using the PYMOL program.

Deletion of the repeat elements in calnexin and calreticulin resulted in 75% of full-length protein binding to $Glc_1Man_9GlcNAc_2$ *in vitro*, indicating that the nonrepeat segments were involved in oligosaccharide binding [11]. A weak affinity for oligosaccharide was found within the repeat segments; however, it displayed a lack of selectivity for monoglucosylated glycans. The electron density map of calnexin soaked in 50-mM glucose demonstrated that the binding site for the terminal glucose of an Asn-linked glycan resides on a concave β-sheet of the globular domain and the glucose lies in proximity to six residues (Y165, K167, Y186, M189, E217, E426) that represent the primary coordination site of the terminal glucose (Figure 11.3) [10]. Corresponding residues in rat calreticulin proved to be crucial for glucose binding, emphasizing the similarity of the lectin sites of calnexin and calreticulin [14].

Calnexin and calreticulin display a high degree of specificity for the glucosidase-trimmed glycan conformation $Glc_1Man_9GlcNAc_2$. The isolated lumenal domain of calnexin is able to selectively bind to the monoglucosylated species in a mixture di-, tri-, and unglucosylated $Man_9GlcNAc_2$ [15].

Calreticulin also has the ability to discriminate between oligosaccharides that differ by a single glucose. In addition, oligosaccharide affinity is diminished on trimming of the glycan from $Glc_1Man_9GlcNAc_2$ to $Glc_1Man_5GlcNAc_2$ and abolished when the α1–6 Man at the B and C branch junction are removed (Figure 11.1) [16]. This branch point mannose has also been found to be essential for calnexin binding [12].

In addition to binding to the terminal glucose, modeling of the calnexin crystal structure predicts that the three A-branch mannoses could bind along the concave lectin site through numerous hydrogen bond contacts with the terminal glucose docked by the six residues of the primary coordination sphere [10]. However, mutation of similar residues in calreticulin predicted to interact with the A-branch mannoses do not prove to be essential for carbohydrate specificity in contrast to the residues in direct contact with the terminal glucose [14]. The high degree of specificity for the monoglucosylated glycan and the structural similarities of the lectin sites in calnexin and calreticulin highlight their importance to the maturation of glycoproteins in the ER.

## B. P-Domain

In the crystal structure of canine calnexin, the P-domain is in an extended hairpin conformation that lies 140 Å away from the globular β-sandwich domain (Figure 11.3) [10]. NMR studies of this isolated region of calreticulin also support the extended hairpin structure of the P-domain [17]. The P-domain contains two repeats (Type I and Type II) that appear four times each in calnexin and three times each in calreticulin [12]. The fold of the extended hairpin results in the repeat motifs overlapping with one another to result in four repeat pairs in calnexin and three pairs in the calreticulin P-domain (Figure 11.2). These repeat regions of calnexin and calreticulin comprise the majority of the sequence similarity among the two chaperones [6].

The Pro-rich P-domain has been found to serve as the docking site for the oxidoreductase ERp57 [11, 18]. The primary site of interaction lies in the region between the repeating elements at the tip of the P-domain hairpin (Figure 11.2) [18]. *In vitro*, deletion mutagenesis of this region resulted in impaired ERp57 binding by both calnexin and calreticulin. However, deletion of some of the surrounding repeat motifs resulted in only slightly impaired ERp57 binding [11]. The residues that comprise the tip of the extended hairpin are primarily negatively charged (Figure 11.2). In calnexin, the residues D342, D344, D346, D348, E350, and E352 have all been found to be important for ERp57 docking. However, calnexin bearing point mutations of single residues was still able to support significant ERp57 binding (Figure 11.3) [19]. Studies on calreticulin have revealed

that residues E239, D241, E243 are the most critical sites for ERp57 docking and have identified an additional amino acid (W244) which may be involved in maintaining the structural organization of these important negatively charged residues [20].

### C. Retention Signals

The differential topologies of calnexin and calreticulin require differing mechanisms for their retention in the ER. ER resident-soluble proteins, including calreticulin, typically contain a C-terminal K-D-E-L motif that triggers the ER retention and retrieval of proteins from the Golgi by the KDEL receptor [21, 22]. The C-terminal KDEL sequence is essential for the retention of calreticulin in the ER [7, 21]. However, calreticulin has also been localized to others locations within the cell including the cell surface and the nucleus [23, 24]. The role of calreticulin in these controversial locations is beyond the scope of this chapter. Generally, type I membrane proteins contain di-Lys motifs on their C-terminal cytosolic tails, which support recycling from the Golgi to the ER via COPI vesicles [25]. However, the C-terminus of calnexin contains an R-K-P-R-R-E sequence that maintains its ER localization [4, 6, 26]. The localization of calnexin and calreticulin is essential for them to perform their functions in protein maturation and quality control.

### D. Regulatory Domains

Both calnexin and calreticulin are $Ca^{2+}$-binding proteins [27, 28]. The crystal structure of canine calnexin identified a $Ca^{2+}$ ion in its globular domain [10]. Association of $Ca^{2+}$ with calnexin and calreticulin is critical for their ability to bind substrates [12, 29]. This has led to the speculation that $Ca^{2+}$ may control the activity of the lectin chaperones. However, it still remains to be demonstrated that the $Ca^{2+}$ levels reached in the ER would be sufficient and altered on a timescale to impact chaperone binding.

The lectin chaperones also display an ability to bind ATP [30, 31]. Binding of nucleotide is thought to result in conformational changes in the chaperones that could in turn regulate their function. Mg-ATP binding to calnexin has been found to cause an increased susceptibility to protease digestion [31]. However, nucleotide binding by the lectin chaperones had no effect on oligosaccharide binding [12]. In addition, the localization of the nucleotide-binding site remains unknown and neither calnexin nor calreticulin displays an ability to hydrolyze ATP [32, 33]. Although both $Ca^{2+}$ and ATP are thought to play a role in regulating calnexin and calreticulin binding to substrate, this issue remains a point of contention.

## IV. The Roles of Calnexin and Calreticulin in Glycoprotein Maturation and Quality Control

Calnexin and calreticulin bind to maturing glycoproteins that have been trimmed sequentially by glucosidase I and II to a monoglucosylated state (Figure 11.4, $Glc_1Man_9GlcNAc_2$) [34, 35]. This binding persists until the final glucose is removed by glucosidase II generating an unglucosylated side chain that is incapable of rebinding to the lectin chaperones. Therefore, glucosidase II serves a dual role in initiating binding as well as inhibiting additional binding. Rebinding to the chaperones is initiated by the UDP-glucose: glycoprotein glucosyltransferase (GT). This protein possesses two key activities, which includes recognition of misfolded substrates and transfer of a single glucose to the unglucosylated A branch. Therefore, binding to the hydrophilic and solute exposed modification is ultimately determined by the GT enzyme through its recognition of exposed hydrophobic surfaces, which are hallmarks for malfolded or unassembled substrates (Figure 11.4).

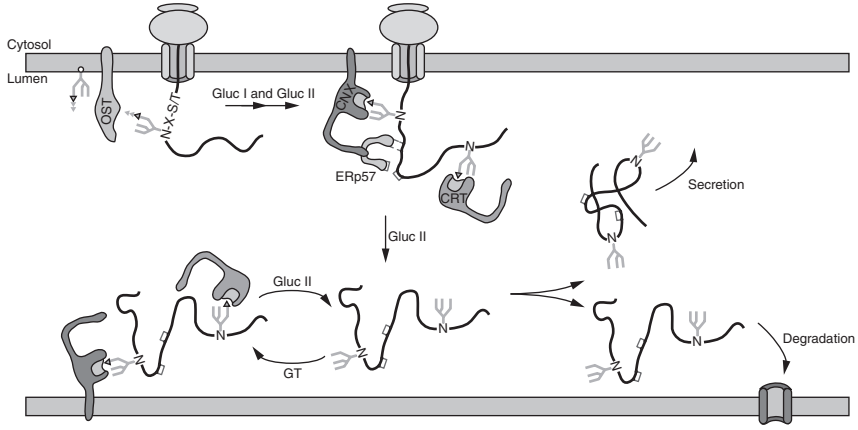



FIG. 11.4. Glycoprotein maturation in the ER. Glycoproteins are cotranslationally translocated into the ER lumen where they are glycosylated by the oligosaccharyl transferase (OST) on the Asn of the consensus sequence Asn-X-Ser/Thr (N-X-S/T). Following glycosylation and glucosidase I and II (Gluc I and II) trimming, the monoglucosylated glycoprotein is bound by the lectin chaperones calnexin (CNX) and calreticulin (CRT), which in turn recruit the oxidoreductase ERp57. The lectin chaperone binding is ablated following the removal of the terminal glucose by glucosidase II activity and the protein is allowed to fold to its native state or deviate toward unproductive aggregates. Proteins bearing exposed hydrophobic regions, a hallmark of misfolding, are reglucosylated by UDP-glucose: glycoprotein glucosyltransferase (GT) and lectin chaperone binding is reinitiated in a process termed the ''calnexin cycle.'' Following rounds of the ''calnexin cycle,'' glycoproteins are either secreted or degraded based on their folding status.

The primary function of the lectin chaperones is to increase the fidelity of glycoprotein maturation by controlling the folding reaction and preventing nonproductive side reactions that lead to protein misfolding and aggregate formation [36, 37]. They do not directly accelerate or catalyze the folding reaction. Rather, persistent binding to the chaperones slows the folding of the substrate [36]. Indeed, in cells deficient in calnexin (*cnx*$^{-/-}$) *or calreticulin* (*crt*$^{-/-}$), the kinetics of glycoprotein folding were accelerated while the fidelity of the folding process was compromised [38–40]. Binding to substrates commences cotranslationally as the protein is being translocated in to the ER lumen, and continues posttranslationally until the properly folded and assembled protein is ready to be packaged into COPII vesicles for anterograde transport to the Golgi. Chaperone binding also serves a quality control function of retaining immature or nonnative substrates in the maturation environment of the ER.

### A. Topological Constraints Control Lectin Chaperone Functions

Calnexin and calreticulin bind a large variety of glycoproteins including soluble, and type I, type II, and polytopic membrane proteins [34, 35, 41–45]. While both calnexin and calreticulin display identical binding specificities for monoglucosylated glycans, they have preferences for glycans situated within different regions of substrate glycoproteins [15, 16, 46]. In general, calnexin binds membrane proximal glycans and calreticulin binds glycans that cannot be reached by calnexin since they emerge deeper into the ER lumen.

Calreticulin but not calnexin posttranslationally binds to the soluble and heavily glycosylated blood coagulation factor V [47]. In contrast, the viral membrane glycoprotein Vesicular stomatitis virus G protein (VSVG) is bound exclusively by calnexin during its maturation [48]. These differences in substrate binding can be attributed to the topology of the glycans on the substrate protein, and the topological location of the binding sites of calnexin and calreticulin within the lumen of the ER. Other glycoproteins such as *influenza* hemagglutinin and tyrosinase associate with both calnexin and calreticulin. When calreticulin was expressed with a membrane anchor in HepG2 cells, its substrate selection became similar to calnexin indicating topology was a key substrate determinant [9, 49]. Conversely, when the membrane anchor was removed from calnexin creating a soluble construct, the profile of substrate binding resembled that of calreticulin [49]. This alludes to a possible lack of redundancy among calnexin and calreticulin in regards to their substrate selection. This is also evident in cells devoid of either calnexin or calreticulin, where the folding of VSVG, which is bound by calnexin exclusively, was unaffected in *crt*$^{-/-}$ cells but impaired in *cnx*$^{-/-}$ cells [40].

However, hemagglutinin folding was much less efficient in cells devoid of calnexin and relatively uncompromised in calreticulin deficient cells, despite interacting with both chaperones during its maturation.

Calnexin and calreticulin initiate their binding to substrates cotranslationally and cotranslocationally [46, 50, 51]. Since the average eukaryotic protein of ~500 amino acids requires about 2 min to be translated, this indicates that the actions of the glucosidases that yield the monoglucosylated substrates are rapid [52]. Ribosome- and translocon-arrested chains lacking stop codons have been utilized to analyze the cotranslational process [53]. These studies have found that calnexin is localized first in the assembly line of lectin chaperones that await the emerging nascent chain [54, 55]. As probed by coimmunoprecipitation, two glycans appear to be required for calnexin binding whereas the subsequent addition of glycans permits cotranslational calreticulin binding. The blood coagulation factor V only binds calreticulin posttranslationally [47]. However, cotranslationally, it was found to initially be bound by calnexin and bind calreticulin after the addition of five glycans [55]. These results are indicative of an organization of the chaperones whereby calnexin is localized nearby the translocon and the soluble calreticulin resides deeper within the ER lumen. Support for this organization has also been provided by studies that show that calnexin can bind ribosomes in a phosphorylation-dependent manner [56]. The phosphorylation of the C-terminal tail by casein kinase 2 and extracellular signal-regulated kinase-1 has been proposed to control the localization of calnexin.

The location and number of the glycans on the nascent chain determine the order of chaperone binding. Proteins that do not contain a glycan in their first 50 amino acids bind initially to the ER Hsp70 family member BiP [51, 55]. BiP binds directly to exposed and extended hydrophobic regions [57]. However, once carbohydrates are added, the lectin chaperone system takes over with calnexin binding the initially added glycans followed by calreticulin [54, 55]. Together the traditional Hsp70 and the carbohydrate chaperone systems protect the vulnerable nascent chain during its early stages of maturation.

It has been proposed that glycans have evolved to be localized to sites on a protein that require protection during the folding process [1, 54]. For instance, hemagglutinin is composed of nonsequential folding domains that result in the formation of a large disulfide loop that covalently links the N- and C-termini of the ectodomain (Cys14–Cys466). Since Cys residues form covalent bonds, they can be problematic residues for protein folding. To ensure that the N-terminal region (Cys14) of hemagglutinin is protected while it waits for the translation of the C-terminal domain, the N-terminus is decorated with multiple conserved Asn-linked glycans (Asn8, Asn22, and

Asn38). This dictates rapid and prolonged binding to calnexin, which holds the membrane proximal N-terminus close to the membrane and inhibits premature oxidation of the N-terminal Cys residue until its native partner Cys466 is translocated. Future studies will be required to explore the magnitude of the involvement of glycans in directing the folding process.

### B. Lectin Chaperones Involvement in Quality Control

In addition to aiding in the proper folding of glycoproteins, calnexin and calreticulin also play central roles in the quality control process. They monitor the maturation of glycoproteins in the ER with the help of GT, assisting initially in retaining immature or aberrant proteins, and then subsequently targeting them for destruction through the ER-associated protein degradation (ERAD) pathway. Following the enzymatic elimination of the terminal α1–3 glucose of the monoglucosylated glycan by glucosidase II, ablating lectin chaperone interaction, the unglucosylated substrate is free to fold to its native state. If an aberrantly folded or assembled structure is detected by the quality control folding sensor GT, the protein will be enzymatically reglucosylated to reinitiate calnexin and calreticulin binding though a process termed the ''calnexin cycle'' [34, 35, 44, 58, 59]. In this cycle, misfolded glycoproteins are redirected to calnexin and calreticulin to promote their folding and support their retention in the ER (Figure 11.4).

Failure to achieve native conformations is the major determinant in entry into the calnexin cycle. The temperature-sensitive variant of VSVG has been shown to have monoglucosylated glycans that are rapidly turned over and reglucosylated [60]. Reglucosylation of VSVG at the nonpermissive temperature reinitiated calnexin binding resulting in ER retention, whereas VSVG at the permissive temperature displayed minimal reglucosylation and was subsequently transported to the Golgi [61]. Expression of a truncated calnexin mutant in COS cells lacking its retention sequence resulted in the exportation of calnexin to the Golgi and the cell surface [26]. This supported the leakage out of the ER of T-cell receptor (TCR) subunits that are normally retained in the ER through calnexin binding. Therefore, calnexin determined the intracellular localization of the unassembled TCR subunits. The importance of the lectin chaperones in the retention of misfolded glycoproteins is most evident in $cnx^{-/-}$ and $crt^{-/-}$ cells, where despite deficiencies in the efficiency of substrate folding when one of the chaperones is absent, the retention of malformed proteins via the lectin chaperones remained unaffected [40].

Terminally misfolded and ER retained proteins are eventually targeted for degradation by the cytosolic proteasome through the ERAD process [1]. This helps to prevent the accumulation of destructive intracellular

aggregates and frees up the chaperones to assist the newly translated nascent chains. Prolonged binding to calnexin has been proposed to lead to ERAD through the extraction of glycoproteins from calnexin binding by ER degradation enhancing α-mannosidase-like protein (EDEM) [62, 63]. EDEM has been shown to be in a complex with calnexin but not with the soluble calreticulin, optimally positioning it for the transfer to EDEM [62]. EDEM is believed to bind to mannose-trimmed substrates, indicative of prolonged ER residency and terminal misfolding. However, many questions still remain on how proteins are sorted and processed by the ERAD machinery. These questions include how EDEM recognizes and treats the aberrant proteins and what is the nature of the translocon that provides the conduit to retrotranslocate the misfolded substrate to the cytosol for proteasomal degradation. Future studies will be required to elucidate the mechanism of protein selection and degradation for ERAD.

### C. Peptide-Binding Function

There are currently two models concerning the mode of lectin chaperone association with protein substrates (Figure 11.5). While the early stages of glycoprotein maturation, such as glycan transfer/trimming and the association of the lectin chaperones with Asn-linked glycans, are essentially the same in either model, differences lie in the general affinity of the lectin chaperones for substrates and the methods by which lectin chaperones bind substrate (Figure 11.4) [8]. The first model, known as the ''lectin-only'' model, hypothesizes that affinity of calnexin and calreticulin for substrate glycoproteins is determined solely by its binding capacity for the monoglucosylated Asn-linked glycan (Figure 11.5, left panel) [34, 35]. The second model is the ''dual-binding'' model, in which calnexin and calreticulin binding would still be driven not only by their affinity for the monoglucosylated glycan but also result from associations with hydrophobic regions in glycoproteins (Figure 11.5, right panel) [15, 32, 37, 64, 65].

The lectin only model suggests that access to the α1–3 glucose, following glucosidase liberation of the two terminal glucoses, is the sole determinant in lectin chaperone binding. Removal of this terminal glucose by glucosidase II ablated any calnexin or calreticulin interaction with *influenza* hemagglutinin [34, 35, 66, 67]. Extensive studies of glycoprotein binding to calnexin and calreticulin using glucosidase inhibitors such as castenospermine (CST) or deoxynojirimycin (DNJ), which prevent glucosidase cleavage of the A-branch glucoses, prohibit the association of the lectin chaperones with substrates such as hemagglutinin, major histocompatibility (MHC) class I heavy chain, HIV gp160, and T-cell antigen receptor proteins [9, 35, 44, 68–70]. Additionally, prevention of glycan transfer by the OST or the enzymatic

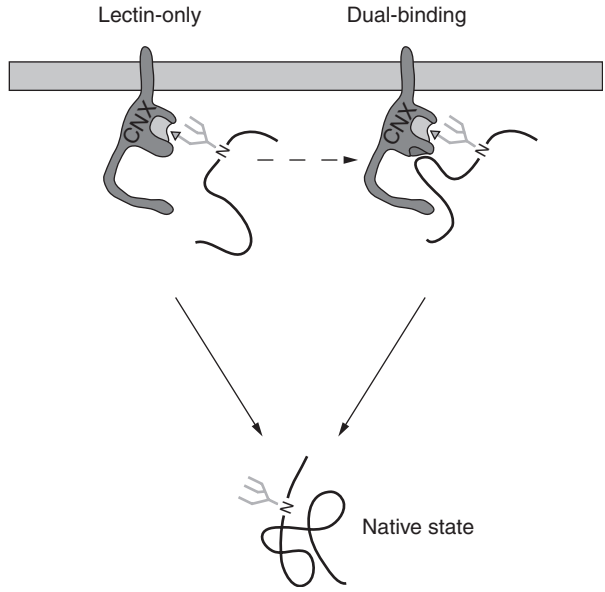



FIG. 11.5. Differing modes of lectin chaperone interaction with glycoprotein substrates. There are two models for the way in which the lectin chaperones bind their substrates. In the ''lectin only'' model the monoglucosylated Asn-linked glycan is the sole requirement for calnexin and calreticulin association with substrate glycoproteins. In the ''dual-binding'' hypothesis, the lectin chaperone–substrate complex formation is initiated by the presence of an Asn-linked glycan and strengthened through an additional polypeptide-binding site on the lectin chaperone. Both models propose that lectin chaperone association promotes the formation of native conformations in substrate glycoproteins.

removal of Asn-linked glycans prevented calnexin and calreticulin substrate interaction [34, 41, 69, 71–73]. Furthermore, in glucosidase-deficient mammalian cells, neither calnexin nor calreticulin was able to associate with endogenously labeled substrate glycoproteins or those introduced by viral infection since the inner most $\alpha$1–3 glucose was unavailable for binding [44, 68, 74]. Binding of the lectin chaperones to unglycosylated glycoproteins that are grossly misfolded has been observed on occasion but this binding has been attributed to massive nonspecific aggregate formation [75].

The dual-binding hypothesis states that calnexin and calreticulin can behave as more canonical chaperones, in that, they can suppress aggregation of vulnerable polypeptides and sustain folding intermediates or malformed proteins in a state in which they can achieve their native conformations [32]. It is believed that the lectin-binding capacity of calnexin and calreticulin initiates the interaction with glycoprotein substrates

and their polypeptide-binding capability serves to strengthen these interactions [32, 65, 76]. This is evident by studies in which glucosidase trimming was inhibited or glycosylation sites on the substrate were mutated and calnexin association was prevented; however, removal of the glycan following calnexin–substrate complex formation failed to result in the release of calnexin from MHC class I heavy chain [65]. Calnexin has been found to be associated with glycoproteins after the enzymatic removal of their glycans in several cases [15, 33, 64] and also under conditions in which glucosidase activity is inhibited [47, 77–79]. In addition, in contrast to previous studies using either glucosidase I or glucosidase II deficient cells [44, 68, 74], calnexin has been observed to associate with endogenously labeled glycoproteins in the same cell lines utilizing a different immunoprecipitation protocol [77].

The lectin chaperones can also suppress the aggregation and promote the folding of natively nonglycosylated purified proteins *in vitro* such as citrate synthase, malate dehydrogenase, and proteolipid protein [32, 33, 76, 80]. Deletion studies have localized the proposed polypeptide-binding site to the globular domain of calnexin and calreticulin and perhaps some regions of the arm or P-domain [11]. Future studies will be necessary to identify the precise location of the peptide-binding site on the calnexin crystal structure and if this polypeptide binding can be regulated by ATP in a manner similar to the classical chaperones.

## V. The Calnexin-Binding Cycle in Yeast

The importance of the calnexin-binding cycle to cellular homeostasis can be explored by analyzing its requirement in lower eukaryotes such as yeast. Calnexin has been identified in both *Schizosaccharomyces pombe* (Cnx1p) and *Saccharomyces cerevisiae* (Cne1p) [81–83]. However, neither *S. pombe* nor *S. cerevisiae* contains a homologue of calreticulin. *S. pombe* calnexin is 39.4% identical to human calnexin and 32.5% identical to that found in *S. cerevisiae*. It is a type I membrane-anchored glycosylated protein and is essential for viability [81, 82]. *S. cerevisiae* calnexin is 24% identical to its canine homologue but differs from *S. pombe* and mammalian calnexin in that it does not contain a C-terminal cytoplasmic region [83]. In addition, Cne1p is not capable of binding to $Ca^{2+}$ and is not essential for viability. The lectin-binding region and the P-domain of Cne1p seemingly function in a similar fashion to its mammalian counterpart [84]. The ER quality control system of *S. pombe* is similar to the system found in the mammalian ER; however, glycoprotein maturation in *S. cerevisiae* is much different due to

the noted differences in calnexin and also the lack of the central folding sensor GT, which supports the calnexin-binding cycle [85, 86].

## VI. ERp57, a Member of the PDI Family of Oxidoreductases

ERp57, also known as ERp60, p58, grp58, and originally as phospholipase Cα, is a component of the ER maturation and quality control system that governs the efficient folding of proteins in the secretory pathway (Figure 11.4). ERp57 appears to act exclusively on glycosylated proteins as dictated by its association with the lectin chaperones calnexin and calreticulin [54, 55, 87–89]. It is a member of the human protein disulfide isomerase (PDI)-like family, which consists of 17 ER members with varying enzymatic properties and functions [90].

The PDI family of oxidoreductases all contain signature thioredoxin folds, first identified in the prokaryotic cytosolic reductase thioredoxin [91]. The most thoroughly studied member of this family is PDI. ERp57 is the most similar family member, both structurally and functionally to PDI, sharing 33% sequence identity [92]. Structural comparisons indicate they are both soluble, contain four distinct thioredoxin-like domains, and share identical organization of their active site Cys residues (C-G-H-C) (Figure 11.6) [90, 91]. Functionally, both proteins are oxidoreductases that perform oxidation, reduction, and isomerization reactions on substrate Cys residues. The active site thioredoxin domains of ERp57 and PDI are denoted **a** and **a′**, whereas catalytically inactive thioredoxin domains lacking the active Cys residues are termed **b** and **b′**. These four domains are arranged in a linear fashion **a-b-b′-a′** (Figure 11.6) [93–95].

### A. Structural Insights into ERp57 from the PDI Crystal Structure

The first crystal structure for full-length PDI has recently been published and it provides important structural information about the yeast PDI, which can also be applied to its closest homologue ERp57 (Figure 11.7) [96]. The structure reveals the orientation of the four domains where the two catalytic domains, **a** and **a′**, are situated on top of the two closely associated noncatalytic domains, **b** and **b′**. Overall, the structures resemble a twisted ''U'' shape (Figure 11.7). The protein–protein interaction interface of **b** and **b′** domains comprises a total surface area of ~700 $Å^2$, which is proposed to form a stable base on which the flexible **a**/**a′** domains can dynamically interact with substrates. Indeed, a surface representation of the PDI

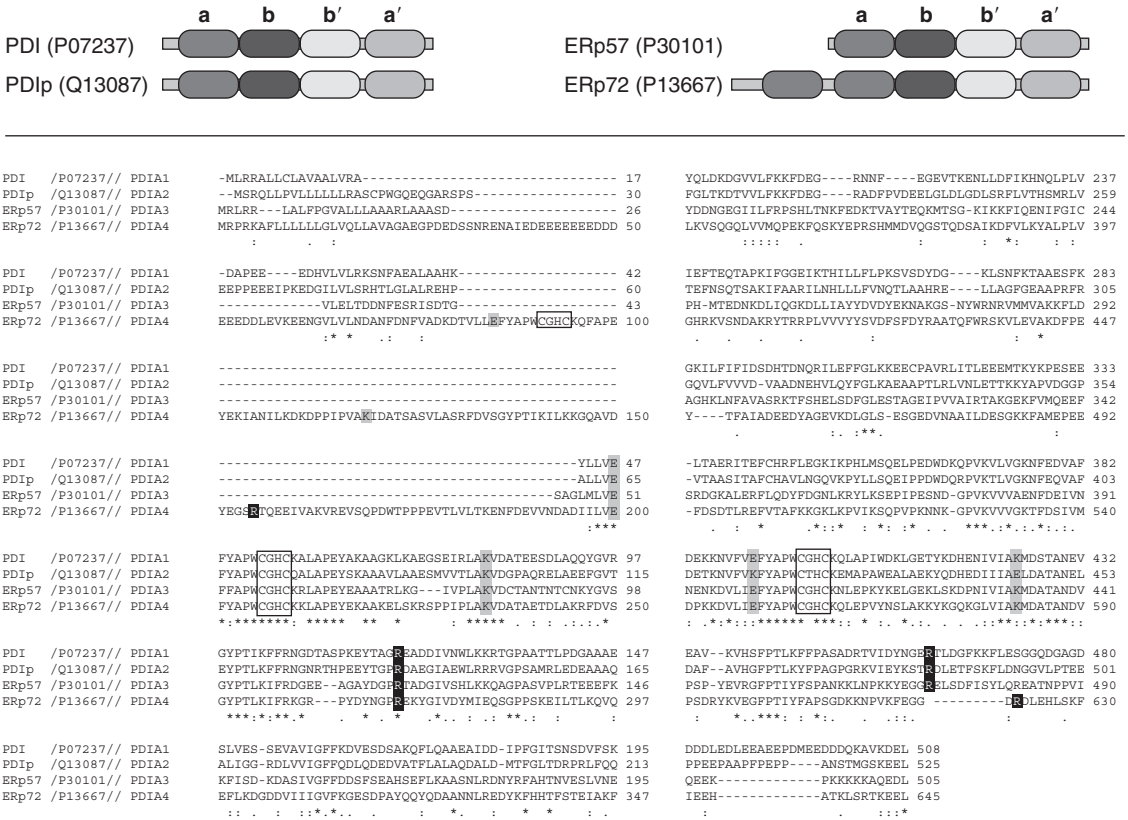



FIG. 11.6. Sequence analysis of PDI family oxidoreductases. Schematics of the thioredoxin domains of four human PDI family members (PDI, PDIp, ERp57, and ERp72) are depicted, with accession numbers indicated. Sequence alignment was performed using the CLUSTALW program [121]. C-G-H-C active sites are marked with a black box. Charged pairs (E and K) and the conserved arginine (R) are shaded gray and black, respectively [90, 103]. The symbols (*), (:), and (.) denote identical, conserved, and semiconserved residues, respectively.

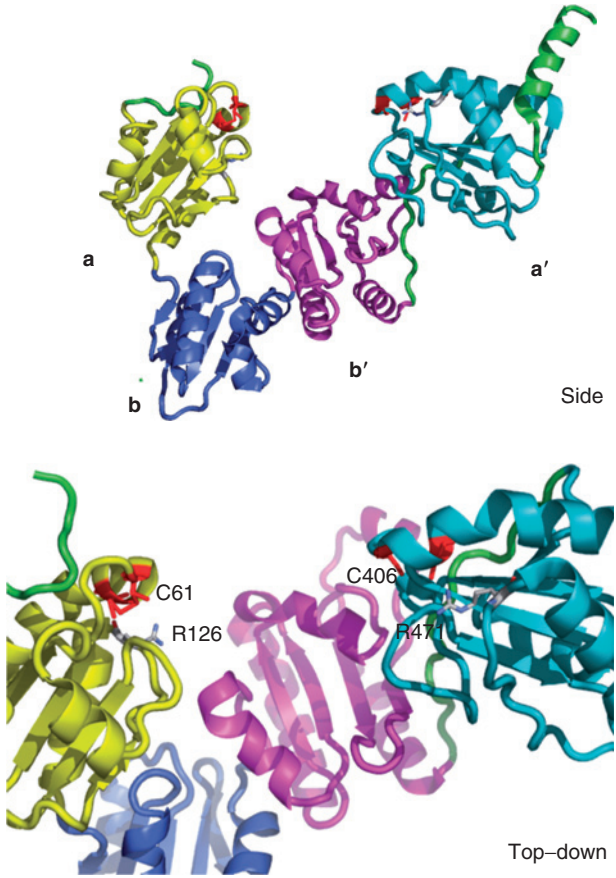



FIG. 11.7. Crystal structure of yeast PDI. The 2.4-Å crystal structure of yeast PDI is depicted in both a side view and a top view (PDB code: 2B5E). The thioredoxin domains **a, b, b′**, and **a′** are colored yellow, blue, purple, and cyan, respectively. Active site Cys residues are colored red and the conserved Arg residue is colored by element, both are represented in stick form and labeled accordingly.

structure reveals a large hydrophobic patch that stretches the length of the cleft formed by four adjoining thioredoxin domains. This may allow freedom in the orientation of the substrate within the cleft so numerous disulfide bonds of varying lengths can be manipulated, and help to funnel substrates to the defined peptide-binding site in the **b′** domain in the center at the bottom of the "U."

Sequence alignment and limited proteolysis demonstrate that ERp57 contains four distinct thioredoxin domains organized in the similar **a-b-b′-a′**

configuration [93, 95, 97]. The hydrodynamic volume of ERp57 indicates that it is an extended conformation similar to the domain architecture of PDI [95]. NMR studies of the isolated **a** and **a′** have revealed thioredoxin folds in both of these domains and a strong resemblance to the equivalent domains of PDI [97, 98]. These similarities among PDI and ERp57 are suggestive of both proteins possessing similar overall structures and functions.

### B. ERp57 Docking onto Its Partner Lectin Chaperones

The lectin chaperone-binding site on ERp57 has been localized to the C-terminus of ERp57 using a yeast ER two-hybrid assay [19]. Further evidence has revealed that the ERp57 **b′** domain is necessary for calreticulin binding; however, it appears that the most efficient binding requires additional contact points from other distal domains [97]. In a detailed mutational study of the **b′** domain, nearly all point mutations disrupted calreticulin docking. Importantly, amino acids critical for PDI substrate binding were also crucial for lectin chaperone binding to ERp57, pointing to a conservation of the **b′** domain among PDI family members. Calnexin and calreticulin seem to act as the substrate for ERp57 binding in a similar fashion to malformed substrates binding to PDI for disulfide formation/rearrangement. All of the residues that were found to be critical for lectin chaperone binding on ERp57 are conserved among all species, underscoring their importance in the function of the oxidoreductase [99].

## VII. Redox Activity of ERp57

ERp57 can function as an oxidase, a reductase, and an isomerase [95]. The redox state of the active site Cys residues determines what type of reactions the protein can participate in. An oxidized active site where a disulfide bond is formed between neighboring Cys in the C-G-H-C motifs can contribute this disulfide bond to the substrate protein resulting in its oxidation (Figure 11.8, oxidation). Here, the active site Cys residues act as an electron acceptor causing their reduction when the oxidation reaction is completed. Alternatively, an ERp57 active site containing reduced Cys residues can support the isomerization or the reduction of substrate disulfides (Figure 11.8). The oxidative state, redox potential, and the activity of the catalytic Cys residues are controlled by the environment of the enzyme, as well as the organization of conserved residues surrounding the active site.

The redox potential of ERp57 lies between that of the canonical reductase thioredoxin and the well-characterized bacterial oxidase DsbA, indicative of its multifunctional redox role (−0.167 V, −0.156 V for the **a** and **a′**

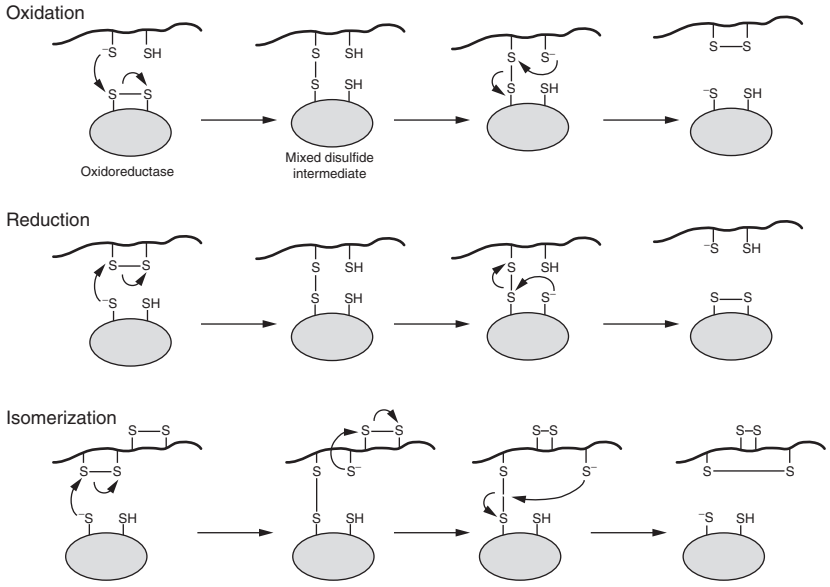



FIG. 11.8. Schematic of the redox reactions of an oxidoreductase. The reaction schemes for an oxidoreductase undergoing oxidative, reductive, and isomerization reactions are indicated. The flow of electrons is marked with arrows. In the case of an oxidative reaction, the thiolate of the substrate begins the nucleophilic attack on the disulfide of the oxidoreductase. Whereas, in reduction or isomerization of the substrate, the thiolate of the oxidoreductase begins the nucelophilic attack on the substrate. All three reactions share the common transition state of a mixed disulfide between enzyme and substrate.

domains, respectively) [95]. The *in vitro* activity of ERp57 in oxidation, reduction, and isomerization reactions was found to be significantly less efficient than PDI. The isolated catalytic domains of ERp57, **a** and **a′**, displayed a decrease in activity when compared to the full-length protein [95, 97].

Mutational studies of the active site residues of ERp57 have shown that by mutating the C-terminal Cys of the Cys-Gly-His-**Cys** motif in both active sites, ERp57 was able to function as a reductase. This indicates that the N-terminal Cys of the **Cys**-Gly-His-Cys motif is the most reactive [100]. However, mutation of the identical residue in the isolated **a′** domain resulted in minimal reductase activity, supporting the notion of the two active sites playing differing roles in redox reactions. Mutations in the central residues of the motif lead to a variety of results indicating that these residues are crucial for the redox activity of the Cys residues. Interestingly, insertion of an additional residue, Cys-Gly-His-Cys to Cys-Gly-His-**Ala**-Cys, created a highly stable disulfide bond that could only be reduced under high

concentrations of reductant and was unable to reform after removal of the reducing agent [101]. This indicates that the spatial arrangement of the active site residues is critical for the efficient oxidoreductase activity of ERp57.

In addition to the important internal residues of the Cys-Gly-His-Cys active site, a charged pair of Glu and Lys residues and an Arg play major roles in regulating the active site chemistry of the catalytic Cys residues. These residues are conserved among many members of the human PDI family, including those that perform oxidoreductase activity (ERp57, ERp72, PDI, PDIp), and have been found to play a role in the redox activities of PDI and ERp57 (Figure 11.6) [90, 93, 102, 103]. The buried Glu maintains a low p$K_a$ for the N-terminal Cys of **Cys**-Gly-His-Cys, and promotes the formation of a reactive thiolate. The Lys residue is situated near the Glu and is responsible for maintaining the Glu in its active state [102, 104]. The positioning of the conserved Arg residue is thought to regulate the p$K_a$ of the C-terminal cysteine of Cys-Gly-His-**Cys**, allowing this residue either to form a reactive thiolate during reduction reactions or to remain protonated during oxidative and isomerization reactions. It has been proposed that this occurs through the movement of the Arg residue away from or toward the catalytic Cys (Figure 11.7) [103]. This mechanism also allows the release of PDI/ERp57 from nonproductive interactions with substrate proteins or in cases where the oxidoreductase is protecting lone Cys residues. The complicated mechanisms of the multifunctional ERp57 highlight its importance to oxidative folding events during glycoprotein maturation.

## VIII. The Role of ERp57 in Glycoprotein Folding

The selection of substrate glycoproteins for ERp57 is driven by its association with the lectin chaperones calnexin and calreticulin. Ternary complexes of calnexin or calreticulin with ERp57 and substrate are required for the manipulation of substrate Cys residues (Figure 11.4). Cross-linking studies have identified a variety of ERp57 glycoprotein substrates, including GLUT1, glycophorin C, preprolactin, and ribonuclease B [87, 88, 105]. ERp57 reaction intermediates have also been trapped directly by the formation of mixed disulfides between ERp57 and the substrate glycoproteins such as Semiliki Forest virus (SFV) viral proteins E1 and p62 (Figure 11.8) [106]. As expected for a foldase, these interactions are transient and lead to the productive folding of the substrate.

The requirement of ERp57 in the oxidative folding of viral glycoproteins was assessed by utilizing an ERp57 knockout lymphocyte cell line

(*ERp57*$^{-/-}$) [107, 108]. ERp57 appeared to be expendable for the cotranslational oxidation of hemagglutinin, a substrate that relies heavily on the lectin chaperone-binding cycle [108]. In contrast, posttranslational oxidative rearrangement of hemagglutinin disulfide bonds resulted in the accumulation of covalently linked aggregates in the absence of ERp57. In addition, there was a dramatic delay in reaching an endo H-resistant state and in the release from calnexin indicating that hemagglutinin folding was compromised. On the basis of these findings, it appears that the critical role of ERp57 in hemagglutinin folding may be as a reductase or an isomerase that acts posttranslationally to reorganize disulfide bonds.

Recent studies have also explored the involvement of ERp57 in the oxidative folding of MHC class I molecules and the formation of the peptide-loading complex prior to its trafficking to the cell surface for antigen presentation. ERp57 associates with the heavy chain of MHC class I during the early stages of its maturation when calnexin is bound [5, 109–111]. Additionally, ERp57 is a component of the later forming peptide-loading complex of MHC class I, consisting of calreticulin, β2-microglobulin, the transporter associated with antigen processing (TAP), the membrane-bound glycoprotein tapasin, and the peptide for extracellular presentation [24, 112]. There is conflicting evidence for the role of ERp57 in the oxidative folding of MHC class I heavy chain prior to its assembly into the peptide-loading complex. Disulfide formation in heavy chains occurred with a decreased efficiency when ERp57 levels were depleted by RNA interference [113]. Oxidative folding was approximately tenfold slower on ER57 knockdown; however, minimal disruption was observed for assembly of the peptide-loading complex components as well as ER to Golgi transport. In contrast, studies with the ERp57$^{-/-}$ cell line found that MHC class I heavy chain folding was unaffected [107]. Formation of the peptide-loading complex occurred; however, its stability was impaired and led to decreased antigen presentation. Further investigations will be required to determine the specific roles of ERp57 in MHC class I maturation and understand why two different methods of depleting ERp57 levels generated differing results.

PDI-like oxidoreductases are thought to have a high degree of substrate specificity with minimal overlap. Interaction with calnexin and calreticulin through the **b′** domain of ERp57 drives its specificity for monoglucosylated substrates [99]. However, recent studies in ERp57$^{-/-}$ cells revealed that the closely related PDI-like oxidoreductase ERp72 could substitute for ERp57 in the posttranslational rearrangement of SFV glycoproteins E1 and p62 [108]. Oxidative folding was found to be unaffected by the absence of ERp57. This observation was attributed to the involvement of ERp72, which was found in mixed disulfides with E1 and p62 in the absence of

ERp57. However, ERp72 associated with these proteins in a glucosidase-independent manner as demonstrated by its continued association in the absence of glucose trimming. Both ERp57 and ERp72 have also been found to associate with the glycoprotein thyroglobulin, which contains 60 disulfide bonds [114, 115]. Additional studies will be needed to more fully understand the potential networks that exist among the ER oxidoreductases and how these networks relate to the cellular stress response pathways and the potential loss of activity of the individual PDI-like family members.

## IX. Regulation of Calcium Signaling

The lectin chaperones and ERp57 system can also work on mature proteins to modulate their activities. Calreticulin and ERp57 regulate the passage and storage of calcium mediated by sarcoendoplasmic reticulum $Ca^{2+}$-ATPase (SERCA2b). SERCA2b is a member of the $Ca^{2+}$-ATPase family, involved in the transport of $Ca^{2+}$ ions across the ER membrane. It contains 11 transmembrane regions and a C-terminal glycosylated domain. In addition, there is a highly conserved disulfide bond in the fourth lumenal loop or L4 [116]. SERCA2b acts in opposition to the $IP_3$ receptor, which triggers $Ca^{2+}$ efflux from the ER lumen on ligand binding.

Previous studies have revealed that coexpression of calreticulin and SERCA2b in *Xenopus laevis* oocytes resulted in decreased influx of $Ca^{2+}$ into the ER lumen [117]. The interaction of calreticulin with SERCA2b is glycan dependent based on glucosidase inhibitor treatment and mutagenesis of the glycosylation consensus sequence [118]. Calnexin was also found to inhibit the activity of the SERCA2b $Ca^{2+}$ pump [119]. The proximity of the conserved Cys residues of SERCA2b to the glycosylation site led to the hypothesis that calreticulin and ERp57 may be involved in $Ca^{2+}$-buffering modulation [118]. Indeed, coexpression of ERp57 with SERCA2b led to decreased pump activity. Additionally, mutation of the conserved Cys residues in SERCA2b resulted in rapid oscillations in $Ca^{2+}$ trafficking. Control of SERCA2b activity by ERp57 occurred in a $Ca^{2+}$ and glycan-dependent fashion and relied on its oxidoreductase activity (Figure 11.9) [120]. All together, the observation that ERp57 regulates the $Ca^{2+}$-buffering capacity of the SERCA2b pump, a mature natively folded protein in the ER membrane, points to a larger role for maintaining the fidelity of secretory protein folding while at the same time regulating $Ca^{2+}$ homeostasis within the cell. This has numerous implications for the regulation of additional activities within the ER.

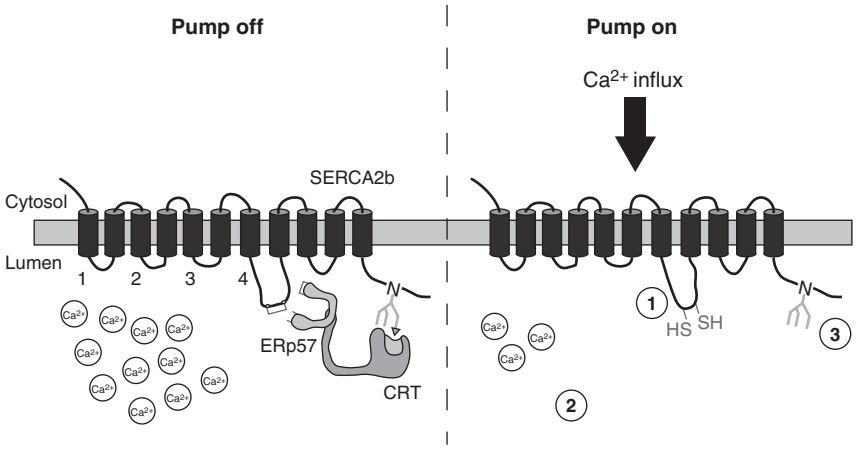



FIG. 11.9. Regulation of the SERCA2b calcium pump by ERp57 activity. When the SERCA2b pump is turned off, ER $Ca^{2+}$ levels are high, the lectin chaperones can associate with a C-terminal SERCA2b glycan, and ERp57 oxidizes a disulfide bond in lumenal loop four (L4). The SERCA2b pump can be switched to the 'on' state under three conditions: (1) ER $Ca^{2+}$ levels are low, which in turn can lead to a decrease in lectin chaperone affinity; (2) since the lectin chaperones cannot bind, ERp57 can no longer oxidize the L4 disulfide bond; and (3) hypothetically, if the C-terminal glycan on SERCA2b were to be deglucosylated then lectin chaperone and ERp57 interaction would be ablated in the absence of fluctuations in the lumenal $Ca^{2+}$ concentration.

## X. Summary

The lectin chaperones and ERp57 play critical roles in maintaining a high fidelity of glycoprotein folding during the early stages of maturation, as proteins emerge from the translocon channel, as well as later stages of protein quality control [1, 2]. Calnexin and calreticulin associate with monoglucosylated proteins, which support the recruitment of their associated oxidoreductase ERp57 to the maturing protein. Calnexin and calreticulin binding can direct the folding of the nascent chain and protect it from nonproductive aggregate formation. ERp57 catalyzes disulfide bond formation, reduction, and isomerization of nascent substrates. This network of chaperones and the oxidoreductase can also modulate the activity of mature proteins, and assist in the presentation of antigen by MHC class I. They are capable of performing these diverse functions because of some fascinating structural characteristics and exquisite active site chemistry. Our understanding of the mechanisms of action of these components has been greatly aided by recently solved crystallographic structures for soluble

calnexin and PDI. Future investigations into these folding factors will help us to better understand the mechanisms of their actions and their implications in maintaining the integrity of the ER quality control system.

## References

1. Hebert, D.N., Garman, S.C., and Molinari, M. (2005). The glycan code of the endoplasmic reticulum: asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol* 15:364–370.
2. Helenius, A., and Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 73:1019–1049.
3. Apweiler, R., Hermjakob, H., and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1473:4–8.
4. David, V., Hochstenbach, F., Rajagopalan, S., and Brenner, M.B. (1993). Interaction with newly synthesized and retained proteins in the endoplasmic reticulum suggests a chaperone function for human integral membrane protein IP90(calnexin). *J Biol Chem* 268:9585–9592.
5. Degen, E., and Williams, D.B. (1991). Participation of a novel 88-kD protein in the biogenesis of murine class I histocompatibility molecules. *J Cell Biol* 112:1099–1115.
6. Wada, I., Rindress, D., Cameron, P.H., Ou, W.-J., Doherty, J.J., II, Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y., and Bergeron, J.J.M. (1991). SSRalpha and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. *J Biol Chem* 266:19599–19610.
7. Fliegel, L., Burns, K., MacLennan, D.H., Reithmeier, R.A., and Michalak, M. (1989). Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum. *J Biol Chem* 264:21522–21528.
8. Williams, D.B. (2006). Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J Cell Sci* 119:615–623.
9. Wada, I., Imai, S.-I., Kai, M., Sakane, F., and Kanoh, H. (1995). Chaperone function of calreticulin when expressed in the endoplasmic reticulum as the membrane-anchored and soluble forms. *J Biol Chem* 270:20298–20304.
10. Schrag, J.D., Bergeron, J.J.M., Li, Y., Borisova, S., Hahn, M., Thomas, D.Y., and Cygler, M. (2001). The structure of calnexin, an ER chaperone involved in quality control of protein folding. *Mol Cell* 8:633–644.
11. Leach, M.R., Cohen-Doyle, M.F., Thomas, D.Y., and Williams, D.B. (2002). Localization of the lectin, ERp57 binding and polypeptide binding sites of calnexin and calreticulin. *J Biol Chem* 277:29686–29697.
12. Vassilakos, A., Michalak, M., Lehrman, M.A., and Williams, D.B. (1998). Oligosaccharide binding characteristics of the molecular chaperone calnexin and calreticulin. *Biochemistry* 37:3480–3490.
13. Hahn, M., Borisova, S., Schrag, J.D., Tessier, D.C., Zapun, A., Tom, R., Kamen, A.A., Bergeron, J.J., Thomas, D.Y., and Cygler, M. (1998). Identification and crystallization of a protease-resistant core of calnexin that retains biological activity. *J Struct Biol* 123:260–264.

14. Kapoor, M., Ellgaard, L., Gopalakrishnapai, J., Schirra, C., Gemma, E., Oscarson, S., Helenius, A., and Surolia, A. (2004). Mutational analysis provides molecular insight into the carbohydrate-binding region of calreticulin: pivotal roles of tyrosine-109 and aspartate-135 in carbohydrate recognition. *Biochemistry* 43:97–106.
15. Ware, F.E., Vassilakos, A., Peterson, P.A., Jackson, M.R., Lehrman, M.A., and Williams, D.B. (1995). The molecular chaperone calnexin binds $Glc_1Man_9GlcNAc_2$ oligosaccharides as an initial step in recognizing unfolded glycoproteins. *J Biol Chem* 270:4697–4704.
16. Spiro, R.G., Zhu, Q., Bhoyroo, V., and Soling, H.D. (1996). Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. *J Biol Chem* 271:11588–11594.
17. Ellgaard, L., Riek, R., Herrmann, T., Guntert, P., Braun, D., Helenius, A., and Wuthrich, K. (2001). NMR structure of the calreticulin P-domain. *Proc Natl Acad Sci* 98:3133–3138.
18. Frickel, E.M., Riek, R., Jelesarov, I., Helenius, A., Wuthrich, K., and Ellgaard, L. (2002). TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain. *Proc Natl Acad Sci USA* 99:1954–1959.
19. Pollock, S., Kozlov, G., Pelletier, M.F., Trempe, J.F., Jansen, G., Sitnikov, D., Bergeron, J.J., Gehring, K., Ekiel, I., and Thomas, D.Y. (2004). Specific interaction of ERp57 and calnexin determined by NMR spectroscopy and an ER two-hybrid system. *EMBO J* 23:1020–1029.
20. Martin, V., Groenendyk, J., Steiner, S.S., Guo, L., Dabrowska, M., Parker, J.M., Muller-Esterl, W., Opas, M., and Michalak, M. (2006). Identification by mutational analysis of amino acid residues essential in the chaperone function of calreticulin. *J Biol Chem* 281:2338–2346.
21. Munro, S., and Pelham, H.R.B. (1987). A C-terminal signal prevents secretion of lumenal proteins. *Cell* 48:899–907.
22. Semenza, J.C., Hardwick, K.G., Dean, N., and Pelham, H.R. (1990). ERD2, a yeast gene required fro the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* 61:1349–1357.
23. Burns, K., Atkinson, E.A., Bleackley, R.C., and Michalak, M. (1994). Calreticulin: from $Ca^{2+}$ binding to control of gene expression. *Trends Cell Biol* 4:152–154.
24. Sadasivan, B., Lehner, P.J., Ortmann, B., Spies, T., and Cresswell, P. (1996). Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity* 5:103–114.
25. Jackson, M.R., Nilsson, T., and Peterson, P.A. (1993). Retrieval of transmembrane proteins to the endoplasmic reticulum. *J Cell Biol* 121:317–333.
26. Rajagopalan, S., Xu, Y., and Brenner, M.B. (1994). Retention of unassembled components of integral membrane proteins by calnexin. *Science* 263:387–390.
27. Baksh, S., and Michalak, M. (1991). Expression of calreticulin in *Escherichia coli* and identification of its $Ca^{2+}$ binding domains. *J Biol Chem* 266:21458–21465.
28. Tjoelker, L.W., Seyfried, C., Eddy, R.L., and Byers, M.G. (1994). Human, mouse, and rat calnexin cDNA cloning: identification of potential calcium binding motifs and gene localization to human chromosome 5. *Biochemistry* 33:3229–3236.
29. Le, A., Steiner, J.L., Ferrell, G.A., Shaker, J.C., and Sifers, R.N. (1994). Association between calnexin and a secretion-incompetent variant of human α1-antitrypsin. *J Biol Chem* 269:7514–7519.
30. Nigam, S.K., Goldberg, A.L., Ho, S., Rohde, M.F., Bush, K.T., and Sherman, M.Y. (1994). A set of endoplasmic reticulum proteins possessing properties of molecular chaperones includes $Ca(2^{+})$-binding proteins and members of the thioredoxin superfamily. *J Biol Chem* 269:1744–1749.

31. Ou, W.J., Bergeron, J.J., Li, Y., Kang, C.Y., and Thomas, D.Y. (1995). Conformational changes induced in the endoplasmic reticulum luminal domain of calnexin by Mg-ATP and $Ca^{2+}$. *J Biol Chem* 270:18051–18059.
32. Ihara, Y., Cohen-Doyle, M.F., Saito, Y., and Williams, D.B. (1999). Calnexin discriminates between protein conformational states and functions as a molecular chaperone *in vitro*. *Mol Cell* 4:331–341.
33. Saito, Y., Ihara, Y., Leach, M.R., Cohen-Doyle, M.F., and Williams, D.B. (1999). Calreticulin functions *in vitro* as a molecular chaperone for both glycosylated and non-glycosylated proteins. *EMBO J* 18:6718–6729.
34. Hammond, C., Braakman, I., and Helenius, A. (1994). Role of N-linked oligosaccharides, glucose trimming and calnexin during glycoprotein folding in the endoplasmic reticulum. *Proc Natl Acad Sci USA* 91:913–917.
35. Hebert, D.N., Foellmer, B., and Helenius, A. (1995). Glucose trimming and reglucosylation determines glycoprotein association with calnexin. *Cell* 81:425–433.
36. Hebert, D.N., Foellmer, B., and Helenius, A. (1996). Calnexin and calreticulin promote folding, delay oligomerization and suppress degradation of influenza hemagglutinin in microsomes. *EMBO J* 15:2961–2968.
37. Vassilakos, A., Cohen-Doyle, M.F., Peterson, P.A., Jackson, M.R., and Williams, D.B. (1996). The molecular chaperone calnexin facilitates folding and assembly of class I histocompatibility molecules. *EMBO J* 15:1495–1506.
38. Howell, D.N., Andreotti, P.E., Dawson, J.R., and Cresswell, P. (1985). Natural killing target antigens as inducers of interferon: studies with an immunoselected, natural killing-resistant human T lymphoblastoid cell line. *J Immunol* 134:971–976.
39. Mesaeli, N., Nakamura, K., Zvaritch, E., Dickie, P., Dziak, E., Krause, K.-H., Opas, M., MacLennan, D.H., and Michalak, M. (1999). Calreticulin is essential for cardiac development. *J Cell Biol* 144:857–868.
40. Molinari, M., Eriksson, K.K., Calanca, V., Galli, C., Cresswell, P., Michalak, M., and Helenius, A. (2004). Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control. *Mol Cell* 13:125–135.
41. Ou, W.J., Cameron, P.H., Thomas, D.Y., and Bergeron, J.J.M. (1993). Association of folding intermediates of glycoproteins with calnexin during protein maturation. *Nature* 364:771–776.
42. Loo, T.W., and Clarke, D.M. (1994). Prolonged association of temperature-sensitive mutants of human P-glycoprotein with calnexin during biogenesis. *J Biol Chem* 269:28683–28689.
43. McGinnes, L.W., and Morrison, T.G. (1995). The role of individual oligosaccharide chains in the activities of the HN glycoprotein of Newcastle disease virus. *Virology* 212:398–410.
44. Peterson, J.R., Ora, A., Nguyen Van, P., and Helenius, A. (1995). Transient, lectin-like association of calreticulin with folding intermediates of cellular and viral glycoproteins. *Mol Biol Cell* 6:1173–1184.
45. Pind, S., Riordian, J.R., and Williams, D.B. (1994). Participation of the endoplasmic reticulum chaperone calnexin(p88, IP90) in the biogenesis of the cystic fibrosis transmembrane conductance regulator. *J Biol Chem* 269:12784–12788.
46. Hebert, D.N., Zhang, J.-X., Chen, W., Foellmer, B., and Helenius, A. (1997). The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. *J Cell Biol* 139:613–623.
47. Pipe, S.W., Morris, J.A., Shah, J., and Kaufman, R.J. (1998). Differential interaction of coagulation factor VIII and factor V with protein chaperones calnexin and calreticulin. *J Biol Chem* 273:8537–8544.

48. Hammond, C., and Helenius, A. (1994). Folding of VSV G protein: sequential interaction with BiP and calnexin. *Science* 266:456–458.
49. Danilczyk, U.G., Cohen-Doyle, M.F., and Williams, D.B. (2000). Functional relationship between calreticulin, calnexin, and the endoplasmic reticulum luminal domain of calnexin. *J Biol Chem* 275:13089–13097.
50. Chen, W., Helenius, J., Braakman, I., and Helenius, A. (1995). Cotranslational folding and calnexin binding of influenza hemagglutinin in the endoplasmic reticulum. *Proc Natl Acad Sci USA* 92:6229–6233.
51. Molinari, M., and Helenius, A. (2000). Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* 288:331–333.
52. Netzer, W.J., and Hartl, F.U. (1997). Recombination of protein domains facilitated by cotranslational folding in eukaryotes. *Nature* 388:343–349.
53. Gilmore, R., Collins, P., Johnson, J., Kellaris, K., and Rapiejko, P. (1991). Transcription of full-length and truncated mRNA transcripts to study protein translocation across the endoplasmic reticulum. *Methods Cell Biol* 34:223–239.
54. Daniels, R., Kurowski, B., Johnson, A.E., and Hebert, D.N. (2003). N-linked glycans direct the cotranslational maturation of influenza hemagglutinin. *Mol Cell* 11:79–90.
55. Wang, N., Daniels, R., and Hebert, D.N. (2005). The cotranslational maturation of the type I membrane glycoprotein tyrosinase: the heat shock protein 70 system hands off to the lectin-based chaperone system. *Mol Biol Cell* 16:3740–3752.
56. Chevet, E., Wong, H.N., Gerber, D., Cochet, C., Fazel, A., Cameron, P.H., Gushue, J.N., Thomas, D.Y., and Bergeron, J.J. (1999). Phosphorylation by CK2 and MAPK enhances calnexin association with ribosomes. *EMBO J* 18:3655–3666.
57. Blond-Elguindi, S., Cwirla, S.E., Dower, W.J., Lipshutz, R.J., Sprang, S.R., Sambrook, J.F., and Gething, M.-J.H. (1993). Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. *Cell* 75:717–728.
58. Sousa, M., and Parodi, A.J. (1995). The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase. *EMBO J* 14:4196–44203.
59. Sousa, M.C., Ferrero-Garcia, M.A., and Parodi, A.J. (1992). Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase. *Biochemitry* 31:97–105.
60. Suh, P., Bergmann, J.E., and Gabel, C.A. (1989). Selective retention of monoglycosylated high mannose oligosaccarides by a class of mutant vesicular stomatitis virus G proteins. *J Cell Biol* 108:811–819.
61. Cannon, K.S., and Helenius, A. (1999). Trimming and readdition of glucose to N-linked oligosaccharides determines calnexin association of a substrate glycoprotein in living cells. *J Biol Chem* 274:7537–7544.
62. Oda, Y., Hosokawa, N., Wada, I., and Nagata, K. (2003). EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. *Science* 299:1394–1397.
63. Molinari, M., Calanca, V., Galli, C., Lucca, P., and Pagnetti, P. (2003). Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. *Science* 299:1397–1400.
64. Arunachalam, B., and Cresswell, P. (1995). Molecular requirements for the interaction of class II major histocompatibility complex molecules and invariant chain with calnexin. *J Biol Chem* 270:2784–2790.
65. Zhang, Q., Tector, M., and Salter, R.D. (1995). Calnexin recognizes carbohydrate and protein determinants of class I major histocompatibility complex molecules. *J Biol Chem* 270:3944–3948.
66. Helenius, A. (1994). How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* 5:253–265.

67. Helenius, A., Trombetta, E.S., Hebert, D.N., and Simons, J.F. (1997). Calnexin, calreticulin and the folding of glycoproteins. *Trends Cell Biol* 7(5):193–200.
68. Balow, J.P., Weissman, J.D., and Kearse, K. (1995). Unique expression of major histocompatibility complex class I proteins in absence if glucose trimming and calnexin association. *J Biol Chem* 270:29025–29029.
69. Otteken, A., and Moss, B. (1996). Calreticulin interacts with newly synthesized human immunodeficiency virus type I envelope glycoprotein, suggesting a chaperone function similar to that of calnexin. *J Biol Chem* 271:97–103.
70. Van Leeuwen, J.E.M., and Kearse, K.P. (1996). The related molecular chaperones calnexin and calreticulin differentially associate with nascent T cell antigen receptor proteins within the endoplasmic reticulum. *J Biol Chem* 272:4179–4186.
71. Nauseef, W.M., McCormick, S.J., and Clark, R.A. (1995). Calreticulin functions as a molecular chaperone in the biosynthesis of myeloperoxidase. *J Biol Chem* 270:4741–4747.
72. Zapun, A., Petrescu, S.M., Rudd, P.M., Dwek, R.A., Thomas, D.Y., and Bergeron, J.J.M. (1997). Conformation independent binding of monoglucoylated ribonuclease B to calnexin. *Cell* 88:29–38.
73. Rodan, A.R., Simons, J.F., Trombetta, E.S., and Helenius, A. (1996). N-linked oligosaccharides are necessary and sufficient for association of RNase B with calnexin and calreticulin. *EMBO J* 15:6921–6930.
74. Ora, A., and Helenius, A. (1995). Calnexin fails to associate with substrate proteins in glucosidase-deficient cell lines. *J Biol Chem* 270:26060–26062.
75. Cannon, K.S., Hebert, D.N., and Helenius, A. (1996). Glycan dependent and independent association of vesicular stomatitis virus G protein with calnexin. *J Biol Chem* 271:14280–14284.
76. Stronge, V.S., Saito, Y., Ihara, Y., and Williams, D.B. (2001). Relationship between calnexin and BiP in suppressing aggregation and promoting refolding of protein and glycoprotein substrates. *J Biol Chem* 276:39779–39787.
77. Danilczyk, U.G., and Williams, D.B. (2001). The lectin chaperone calnexin utilizes polypeptide-based interactions to associate with many of its substrates *in vivo*. *J Biol Chem* 276:25532–25540.
78. Van Leeuwen, J.E.M., and Kearse, K.P. (1996). Calnexin associates exclusively with individual CD3delta and T cell antigen receptor (TCR) alfa proteins containing incompletely trimmed glycans that are not assembled into multisubunit TCR complexes. *J Biol Chem* 271:9660–9665.
79. Zhang, Q., and Salter, R.D. (1998). Distinct patterns of folding and interactions with calnexin and calreticulin in human class I MHC proteins with altered N-glycosylation. *J Immunol* 160:831–837.
80. Swanton, E., High, S., and Woodman, P. (2003). Role of calnexin in the glycan-independent quality control of proteolipid protein. *EMBO J* 22:2948–2958.
81. Jannatipour, M., and Rokeach, L.A. (1995). The *schitzosaccharomyces pombe* homologue of the chaperone calnexin is essential for viability. *J Biol Chem* 270:4845–4853.
82. Parlati, F., Dignard, D., Bergeron, J.J., and Thomas, D.Y. (1995). The calnexin homologue cnx1$^{+}$ in *Schizosaccharomyces pombe*, is an essential gene which can be complemented by its soluble ER domain. *EMBO J* 14:3064–3072.
83. Parlati, F., Dominguez, M., Bergeron, J.J., and Thomas, D.Y. (1995). *Saccharomyces cerevisiae* CNE1 encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. *J Biol Chem* 270:244–253.

84. Xu, X., Azakami, H., and Kato, A. (2004). P-domain and lectin site are involved in the chaperone function of *Saccharomyces cerevisiae* calnexin homologue. *FEBS Lett* 570:155–160.
85. Fernandez, F.S., Trombetta, S.E., Hellman, U., and Parodi, A.J. (1994). Purification to homogeneity of UDP-glucose: glycoprotein glucosyltransferase from *Schizosaccharomyces pombe* and apparent absence of the enzyme from *Saccharomyces cerevisae*. *J Biol Chem* 269:30701–30706.
86. Parodi, A.J. (1999). Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochim Biophys Acta* 1426:287–295.
87. Elliott, J.G., Oliver, J.D., and High, S. (1997). The thiol-dependent reductase ERp57 interacts specifically with N-glycosylated integral membrane proteins. *J Biol Chem* 272:13849–13855.
88. Oliver, J.D., van der Wal, F.J., Bulleid, N.J., and High, S. (1997). Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* 275:86–88.
89. Van der Wal, F.J., Oliver, J.D., and High, S. (1998). The transient association of ERp57 with N-glycosylated proteins is regulated by glucose trimming. *Eur J Biochem* 256:51–59.
90. Ellgaard, L., and Ruddock, L.W. (2005). The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep* 6:28–32.
91. Freedman, R.B., Hirst, T.R., and Tuite, M.F. (1994). Protein disulphide isomerase: building bridges in protein folding. *Trends Biochem Sci* 19:331–336.
92. Ferrari, D.M., and Soling, H.D. (1999). The protein disulphide-isomerase family: unravelling a string of folds. *Biochem J* 339(Pt. 1):1–10.
93. Alanen, H.I., Salo, K.E., Pekkala, M., Siekkinen, H.M., Pirneskoski, A., and Ruddock, L.W. (2003). Defining the domain boundaries of the human protein disulfide isomerases. *Antioxid Redox Signal* 5:367–374.
94. Alanen, H.I., Williamson, R.A., Howard, M.J., Lappi, A.K., Jantti, H.P., Rautio, S.M., Kellokumpu, S., and Ruddock, L.W. (2003). Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member. *J Biol Chem* 278:28912–28920.
95. Frickel, E.M., Frei, P., Bouvier, M., Stafford, W.F., Helenius, A., Glockshuber, R., and Ellgaard, L. (2004). ERp57 is a multifunctional thiol-disulfide oxidoreductase. *J Biol Chem* 279:18277–18287.
96. Tian, G., Xiang, S., Noiva, R., Lennarz, W.J., and Schindelin, H. (2006). The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites. *Cell* 124:61–73.
97. Silvennoinen, L., Myllyharju, J., Ruoppolo, M., Orru, S., Caterino, M., Kivirikko, K.I., and Koivunen, P. (2004). Identification and characterization of structural domains of human ERp57: association with calreticulin requires several domains. *J Biol Chem* 279:13607–13615.
98. Silvennoinen, L., Karvonen, P., Koivunen, P., Myllyharju, J., Kivirikko, K., and Kilpelainen, I. (2001). Assignment of 1H, 13C and 15N resonances of the a′ domain of ERp57. *J Biomol NMR* 20:385–386.
99. Russell, S.J., Ruddock, L.W., Salo, K.E., Oliver, J.D., Roebuck, Q.P., Llewellyn, D.H., Roderick, H.L., Koivunen, P., Myllyharju, J., and High, S. (2004). The primary substrate binding site in the b′ domain of ERp57 is adapted for endoplasmic reticulum lectin association. *J Biol Chem* 279:18861–18869.
100. Antoniou, A.N., Ford, S., Alphey, M., Osborne, A., Elliot, T., and Powis, S.J. (2002). The oxidoreductase ERp57 efficiently reduces partially folded in preference to fully folded MHC class I molecules. *EMBO J* 21:2655–2663.

101. Beynon-Jones, S.M., Antoniou, A.N., and Powis, S.J. (2006). Mutational analysis of the oxidoreductase ERp57 reveals the importance of the two central residues in the redox motif. *FEBS Lett* 580:1897–1902.
102. Chivers, P.T., and Raines, R.T. (1997). General acid/base catalysis in the active site of *Escherichia coli* thioredoxin. *Biochemistry* 36:15810–15816.
103. Lappi, A.K., Lensink, M.F., Alanen, H.I., Salo, K.E., Lobell, M., Juffer, A.H., and Ruddock, L.W. (2004). A conserved arginine plays a role in the catalytic cycle of the protein disulphide isomerases. *J Mol Biol* 335:283–295.
104. Dyson, H.J., Jeng, M.F., Tennant, L.L., Slaby, I., Lindell, M., Cui, D.S., Kuprin, S., and Holmgren, A. (1997). Effects of buried charged groups on cysteine thiol ionization and reactivity in *Escherichia coli* thioredoxin: structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry* 36:2622–2636.
105. Zapun, A., Darby, N.J., Tessier, D.C., Michalak, M., Bergeron, J.J., and Thomas, D.Y. (1998). Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J Biol Chem* 273:6009–6012.
106. Molinari, M., and Helenius, A. (1999). Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells [see comments]. *Nature* 402:90–93.
107. Garbi, N., Tanaka, S., Momburg, F., and Hammerling, G.J. (2006). Impaired assembly of the major histocompatibility complex class I peptide-loading complex in mice deficient in the oxidoreductase ERp57. *Nat Immunol* 7:93–102.
108. Solda, T., Garbi, N., Hammerling, G.J., and Molinari, M. (2006). Consequences of ERp57 deletion on oxidative folding of obligate and facultative clients of the calnexin cycle. *J Biol Chem* 281:6219–6226.
109. Farmery, M.R., Allen, S., Allen, A.J., and Bulleid, N.J. (2000). The role of ERp57 in disulfide bond formation during the assembly of major histocompatibility complex class I in a synchronized semipermeabilized cell translation system. *J Biol Chem* 275:14933–14938.
110. Lindquist, J.A., Jensen, O.N., Mann, M., and Hammerling, G.J. (1998). ER-60, a chaperone with thiol-dependent reductase activity involved in MHC class I assembly. *EMBO J* 17:2186–2195.
111. Morrice, N.A., and Powis, S.J. (1998). A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules. *Curr Biol* 8:713–716.
112. Hughes, E.A., and Cresswell, P. (1998). The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex. *Curr Biol* 8:709–712.
113. Zhang, Y., Baig, E., and Williams, D.B. (2006). Functions of ERp57 in the folding and assembly of major histocompatibility complex class I molecules. *J Biol Chem* 281:14622–14631.
114. Kuznetsov, G., Chen, L.B., and Nigam, S.K. (1994). Several endoplasmic reticulum stress proteins, including ERp72, interact with thyroglobulin during its maturation. *J Biol Chem* 269(37):22990–22995.
115. Di Jeso, B., Park, Y.N., Ulianich, L., Treglia, A.S., Urbanas, M.L., High, S., and Arvan, P. (2005). Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase. *Mol Cell Biol* 25:9793–9805.
116. Bayle, D., Weeks, D., and Sachs, G. (1995). The membrane topology of the rat sarcoplasmic and endoplasmic reticulum calcium ATPases by *in vitro* translation scanning. *J Biol Chem* 270:25678–25684.
117. Camacho, P., and Lechleiter, J.D. (1995). Calreticulin inhibits repetitive intracellular $Ca^{2+}$ waves. *Cell* 82:765–771.
118. John, L.M., Lechleiter, J.D., and Camacho, P. (1998). Differential modulation of SERCA2 isoforms by calreticulin. *J Cell Biol* 142:963–973.

119. Roderick, H.L., Lechleiter, J.D., and Camacho, P. (2000). Cytosolic phosphorylation of calnexin controls intracellular Ca(2+) oscillations via an interaction with SERCA2b. *J Cell Biol* 149:1235–1248.
120. Li, Y., and Camacho, P. (2004). $Ca^{2+}$-dependent redox modulation of SERCA 2b by ERp57. *J Cell Biol* 164:35–46.
121. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31:3497–3500.

This page intentionally left blank

# Part III

## *Crossing Mitochondrial Membranes*

This page intentionally left blank

# 12

# *TOM and SAM Machineries in Mitochondrial Protein Import and Outer Membrane Biogenesis*

MICHAEL JAMES DAGLEY[a,b] • TREVOR LITHGOW[a,b]

[a]*Department of Biochemistry and Molecular Biology*
*The University of Melbourne*
*Parkville, VIC 3010, Australia*

[b]*Bio21 Molecular Science and Biotechnology Institute*
*30 Flemington Road, Parkville, VIC 3010, Australia*

## I. Abstract

Mitochondria are essential organelles. As the mitochondrial genome encodes only a few proteins, the vast majority of proteins that contribute to mitochondrial function are translated in the cytosol from nuclear-encoded transcripts and must be selectively imported into the organelle. The specificity of protein import is achieved by the translocase of the outer mitochondrial membrane (TOM) complex: a multisubunit receptor complex and translocation pore. Further molecular machines including the sorting and assembly machinery (SAM) complex of the mitochondrial outer membrane then transfer and assemble proteins into one of the four mitochondrial subcompartments.

Both the TOM and SAM complexes exhibit a modular structure. The core TOM complex is composed from the most highly conserved subunits, and is complemented by more recently evolved add-on modules, such as the receptors Tom20 and Tom70. These new modules provide sophistication to



DOI: 10.1016/S1874-6047(07)25012-7

the complex, enhancing the specificity and capacity for binding substrates. Likewise, a core SAM complex is critical for the assembly of β-barrel proteins into the outer mitochondrial membrane. Additional modules such as the metaxin-type proteins Sam37 and Sam35 and the yeast proteins Mim1 and Mdm10 are required for the assembly of some, but not all, SAM substrates. Mdm10 is an excellent example of a modular component, being found in at least one other complex where it assists with mitochondrial distribution and morphology. The concept of modular design provides an interesting new perspective on functional aspects of the protein import machinery in mitochondria.

## II. Introduction

Mitochondria are essential organelles found in all eukaryotes. They host a range of metabolic reactions [1], participating in β-oxidation of fatty acids [2], the biosynthesis of amino acids [3] and iron-sulfur clusters [4], and at least in animals, in the process of programmed cell death [5, 6]. For many organisms living in aerobic conditions, mitochondria provide the bulk of cellular ATP production through oxidative phosphorylation [7, 8].

The basic functions of mitochondria are closely linked to their evolutionary origins. These organelles arose as a result of an endosymbiotic event in which a bacterium was taken up by a proto-eukaryote and maintained due to mutual benefit [9]. Over time, most genes from the proto-mitochondrion were lost from the mitochondrial genome to the nuclear genome, creating a unique problem: the mitochondrial proteins encoded in the nuclear genome and synthesized in the cytosol must be targeted to mitochondria, imported and sorted to their final submitochondrial destination. To this end, a number of molecular machines have evolved. These machines are composed of proteins that originated from the bacterial endosymbiont itself, as well as more recently contrived molecules originating from the host organism [10]. In the outer membrane of mitochondria, these machines are the translocase of the outer mitochondrial membrane (TOM) complex (reviewed in [11, 12]), the initial point of entry and sorting for all cytosol-synthesized mitochondria proteins, and the sorting and assembly machinery (SAM) complex that assembles outer membrane proteins with complicated topologies (reviewed in [13–15]) (Figure 12.1). The translocases of the inner mitochondrial membrane (TIM) complexes are also present, reviewed in [16, 17], and covered in Chapters 14 and 15. Following recognition, substrate proteins are translocated and sorted by the TOM complex to one of the four mitochondrial compartments: the outer membrane, the intermembrane space, the inner membrane, and the matrix. Outer membrane

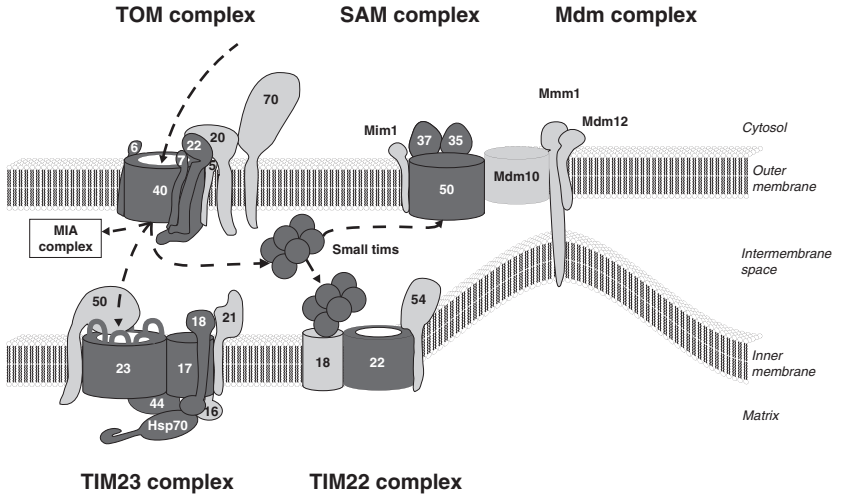



FIG. 12.1. The protein translocation machinery of the outer and inner mitochondrial membranes. Highly conserved components of the core TOM and SAM complexes, the small tims, and the inner membrane TIM23 and TIM22 complexes are shaded dark gray. Additional modules implicated in protein import and assembly are shaded light gray. Arrows indicate the direction of substrate flow between the complexes.

proteins are inserted in a process requiring the TOM complex. Multitopic inner membrane proteins are inserted via the TIM22 complex, matrix proteins are inserted via the TIM23 complex [17], and intermembrane space proteins are folded and assembled by a disulfide relay system that includes Mia40 and Erv1 [18]. In order to accomplish this demanding task, a range of targeting signals have evolved to correctly address incoming proteins to their required destinations.

## A. STRUCTURAL FEATURES OF MITOCHONDRIAL PROTEINS

The mitochondrial import signal is not defined by a primary sequence motif. Instead, mitochondrial proteins have several structural features that designate them and enable their recognition by the protein translocases. These mitochondrial targeting features, which include an overall ''high'' level of hydrophobicity and segments of amphipathic helix, have evolved from the preexisting properties of the bacterial proteins themselves: it has been demonstrated that ~5% of extant bacterial proteins from *Escherichia coli* would target to mitochondria [19]. Furthermore, a percentage of random sequences generated from *E. coli* DNA (2.7%) or the mouse cytosolic protein dihydrofolate reductase (5%) were able to restore mitochondrial

localization to a specific protein lacking its own targeting sequence [20]. This is evidence that small gene rearrangements through the course of evolution would readily direct proteins to the mitochondria [20]. The conserved properties required for mitochondrial targeting are understood, and this knowledge has been employed in the development of a number of tools, such as MITOPRED [21] and PREDOTAR [22], for predicting mitochondrial localization from a primary amino acid sequence.

Most mitochondrial proteins are more hydrophobic than proteins that remain in the cytosol [23], and this feature appears to impede the folding of mitochondrial precursor proteins after translation on cytosolic ribosomes. This ''unfolded'' feature has two consequences that serve to assist targeting to mitochondria: first, any basic, amphipathic targeting sequence will be displayed openly to the protein translocases and, second, the precursor protein will be presented at the mitochondrial surface with attendant chaperones [24–29]. Since the chaperones can dock to the TOM complex [30, 31] (reviewed in [32, 33]), this assists delivery of protein substrates. Whether assisted by chaperones, a critical targeting stage requires the binding of a basic, amphipathic segment of the substrate protein with components of the TOM complex.

In the simplest case, an N-terminal helical segment with basic, amphipathic character can be necessary and sufficient for targeting to mitochondria. Common substrates used in mitochondrial import studies are the artificial precursors CoxIV-DHFR and Su9-DHFR. These proteins are fusions of theN-terminal 14 or 96 residues of cytochrome *c* oxidase subunit IV or $F_0$-ATPase subunit 9, respectively, to the cytosolic protein dihydrofolate reductase. These fusion proteins are imported into mitochondria when expressed *in vivo*, or presented *in vitro* to isolated organelles. However many, perhaps a majority, of the mitochondrial proteins do not have N-terminal sequences that function in targeting, but rather have internal features that serve as targeting signals. In all cases, it appears that these sequences share amphipathic properties and can form $\alpha$-helices to display positively charged residues that assist in the recognition of the substrate as ''mitochondrial'' at acidic surfaces on the protein import complex. The charge characteristics of the substrates and their receptors led to the proposal that an ''acid chain'' guides substrate along the import complex to the translocation channel [34, 35].

In many membrane proteins, mitochondrial targeting information is conveyed in the amphipathic nature of their transmembrane segments. This includes a range of proteins in the inner membrane such as the carrier proteins and components of the electron transport chain complexes, the yeast outer membrane such as Fis1 and OM45, and the Bcl-2 apoptotic pathway proteins in animals [36–38]. The difference between membrane proteins targeted to the mitochondria and those targeted to the endoplasmic

reticulum is that the mitochondrial transmembrane segments are more amphipathic [39–41]. While different mitochondrial targeting sequences are used to direct proteins to one of the four possible mitochondrial compartments, in all cases the relatively abundant TOM complex is responsible for the recognition and sorting of these molecules [42–45].

## III. The TOM Complex

The TOM complex is a multisubunit hetero-oligomeric complex [46]. In yeast, the complex is composed of eight subunits organized in distinct modules (Figure 12.1). The core complex serves as the translocation channel through the outer membrane, consisting of a β-barrel, Tom40, associated with a series of tail-anchored proteins (Tom22, Tom5, Tom6, and Tom7) [36, 47]. This core complex is highly stable, resistant to treatment with 4 M urea and detergents [48, 49]. Interestingly, the most evolutionarily conserved TOM components show the greatest resistance to urea treatment: Tom22 and Tom7 are the two proteins most strongly associated with yeast Tom40 [48]. Tom5, an additional subunit so far only identified in yeast, is coprecipitated as a component of the core complex [50], though is less tightly bound than Tom6 and Tom7 [48]. The highly conserved common core architecture of the complex is consistent with its early evolution, as predicted by its essential and fundamental function. In yeast, Tom20, Tom70, and Tom71 are additional subunits that serve as receptors for the TOM complex. Under optimal solubilization conditions, the entire holocomplex can be purified, such that Tom20 and Tom70 copurify in approximately stoichiometric amounts with the core complex [51]. These additional subunits can be considered modules of the TOM complex that have been added on over time: they appear to have been derived sometime after the divergence of the animal and fungal lineages from other eukaryotes [52, 53].

### A. The Core TOM Complex: A Primitive Protein Translocase?

#### *1. Tom40*

The channel through which unfolded precursor proteins [54] traverse the outer membrane is formed by Tom40. This was first demonstrated by trapping a precursor protein in transit in the ''import site'' of the outer membrane and photo-cross-linking it to its nearest neighbor, a 42-kDa protein referred to as Isp42 and since renamed Tom40 [55]. Tom40 is essential for viability [56] and is predicted to form a β-barrel structure

from 16 anti-parallel β-strands [57, 58]. The Tom40 channel has a hydrophilic pore ~22 Å in diameter, sufficient for the translocation of unfolded proteins. Independent experimental approaches using isolated holo-TOM complex, core TOM complex, and refolded recombinant Tom40 show agreement on the pore diameter; the various techniques used include electron microscopy [51], electrophysiology [59], and the import of precursor protein fused to gold nanoparticles [60]. By observing the effects of antibody binding to the isolated TOM complex, it was demonstrated that the N- and C-termini of Tom40 are located on the same side of the membrane, with protease shaving experiments demonstrating that the termini extend into the intermembrane space [61].

The function of Tom40 extends beyond merely providing a means for unfolded substrates to traverse the membrane. The Tom40 channel is in fact a sorting station, directing substrates to the outer membrane, the intermembrane space, the TIM23 complex, or the TIM22 complex [62]. These sorting pathways are distinct, with a mutation identified which can block one of these paths selectively [62]. This sorting role suggests that the interior of the channel must interact specifically with substrates. Indeed, the channel is not a nonstick pore but has a chaperone-like function: the translocation channel interacts directly with unfolded precursors through hydrophobic surfaces [63]. Furthermore, while the import of signal-anchor proteins (such as Tom20 and OM45) into the outer mitochondrial membrane is also mediated by Tom40, this occurs independently of the channel [64]. It has been proposed that the outer surfaces of Tom40 and some of the transmembrane segments of other TOM components may provide a point of entry into the membrane for these molecules [65].

Studies of recombinant Tom40 from *Neurospora crassa* and *Saccharomyces cerevisiae* compared to isolated TOM complexes show that Tom40 can be refolded into a protein of high β-sheet content which displays electrophysiological properties very similar to the isolated TOM complexes [66]. Electrophysiology studies have revealed that the TOM complex channel is voltage-gated and cation selective, and suggest that it is present as a homodimer [61]. A key finding of these experiments is that the yeast core TOM complex exhibits two coupled channel pores, while the isolated Tom40 exhibits single-channel behavior [66]. Furthermore, when mitochondrial presequences are introduced to either side of isolated TOM complex, a specific change in conductance can be observed [61]. However, as there is little or no transmembrane potential across the outer membrane *in vivo*, it has been proposed that the channel is acting not as a voltage-gated channel, but as a presequence-gated channel [61]. This presequence binding will result in conformational changes to the complex [61, 67].

An early model for the stoichiometry of the TOM complex suggested Tom40, Tom22, Tom70, and Tom20 exist in ratios of 8 (±0.3):3.1 (±0.1):1.5 (±0.2):2 [51]. This stoichiometry is calculated from phosphorimage analysis of radiolabeled proteins obtained from affinity-purified TOM complex. However, these ratios are viewed with some caution given that the calculated eight Tom40 molecules per TOM complex would require two or three protomers per channel to produce the multichannel forms predicted by electrophysiology and visualized by electron microscopy. The precise number of Tom40 channels that are present in the native TOM complex is not clear, with various electron microscopy studies of the TOM complex revealing the presence of one-, two-, and three-pore structures [51, 68, 69]. While there is no general consensus as to the number of Tom40 pores present in the holo-TOM complex, a working model suggests that the TOM complex is a three-pore structure, but that the loss of Tom20 causes the complex to adopt a two-pore structure [68]. When isolated from Δ*tom20* yeast, the TOM complex appears ~50 kDa smaller than the wild-type holo-TOM complex by blue-native PAGE, and reveals only two-ring structures by electron microscopy [68]. Using a different solubilization protocol, a TOM complex lacking Tom70 (but including Tom20) was purified from wild-type yeast, forming a complex of ~440 kDa (a similar size to wild-type holo-TOM complexes) on blue-native PAGE [68]. Electron microscopy revealed this to be a three-ring complex [68]. Crystallization of the core and holo-TOM complexes is the obvious next step to determine both the stoichiometry and spatial arrangement of the components within the machine.

The sequence conservation seen in Tom40 is striking, with highly conserved plant, yeast, and mammalian [70] homologues. Evidence of this conservation is the fact that the *N. crassa* Tom40 can be successfully imported into mitochondria from the protozoan *Trypanosoma brucei*, where it assembles into an ~370-kDa complex [71]. This reveals that not only can the membrane machinery of *T. brucei* recognize elements of the *N. crassa* Tom40 sequence to allow its import, but that an attempt is made to assemble this protein into an oligomeric form. The complex formed from *N. crassa* Tom40 in *T. brucei* mitochondria includes an as-yet unidentified 16-kDa component [71]. It is possible that this protein may be a SAM complex molecule acting to assemble and integrate this foreign β-barrel protein (such as a Mim1 homologue), or a genuine *T. brucei* TOM component which is capable of docking with the fungal Tom40. The latter would be most remarkable given the number of specific protein interactions Tom40 appears to make with other TOM components in order to assemble and function.

*2. Tom22*

Tom22 is an ~17-kDa subunit of the TOM complex [72]. It is anchored in the outer membrane via a central transmembrane segment (Figure 12.2A), thereby exposing an N-terminal domain to the cytosol and a 33-residue domain in the intermembrane space [73]. In yeast, both extramembrane domains of Tom22 are highly acidic, and therefore capable of interacting with presequence-bearing proteins [74].

The acidic N-terminal receptor domain of Tom22 is likely to be poorly structured, with a remarkable amino acid composition. In yeast, the N-terminal receptor domain of Tom22 has 10.3% E, 14.4% D, 8.2% S, 6.2% T, and 4.1% Q: with these five residues being those enriched in disordered regions of other proteins [75, 76]. Using the IUPred tool for the prediction of intrinsically unstructured regions within a protein [77], all known Tom22 sequences exhibit this characteristic at both N- and C-terminal regions (Lithgow, unpublished data). One prospect is that only upon the binding of substrate proteins is order induced in the Tom22 cytosolic domain. Tom22 can be cross-linked to substrate proteins and can function as a receptor directly binding to substrate protein targeting sequences [34]. The acidic character is not an essential feature for the overall function of Tom22 in the TOM complex since extensive mutagenesis of the acidic residues in the fungal Tom22 does not inhibit protein import into mitochondria [78], and in plants and protozoans the acidic part of the N-terminal domain is missing altogether, with a smaller 8-kDa form of Tom22 (Tom22′) found in the TOM complex of these eukaryotes [58, 79, 80]. Still, a receptor domain is required since truncation of the cytosolic domain of Tom22 results in a protein which is correctly located in mitochondria but cannot assist the import of proteins, resulting in a lethal phenotype [81].

The human Tom22 has a very low-sequence similarity to the *N. crassa* sequence, but exhibits a remarkably similar hydrophobicity profile and (negative) charge characteristics in the N-terminus, and the human protein can complement for the loss of *N. crassa* Tom22 [82, 83]. In considering the Tom22 sequences detected in all manner of eukaryotes, the most highly conserved feature is its transmembrane segment (Figure 12.2A). Conserved within this transmembrane stretch are many residues, such as serine, threonine, glycine, and proline that are atypically found within bilayer-spanning α-helices [58, 84, 85]. Given the conservation in the spacing of these residues relative to the water–lipid interface, it seems likely that they would assist protein–protein, rather than protein–lipid, interactions. It is through its transmembrane segment that Tom22 binds tightly to the Tom40 channel and assists in the organization of the TOM complex [86]; the conserved residues in Figure 12.2A represent at least some of the likely points of contact.

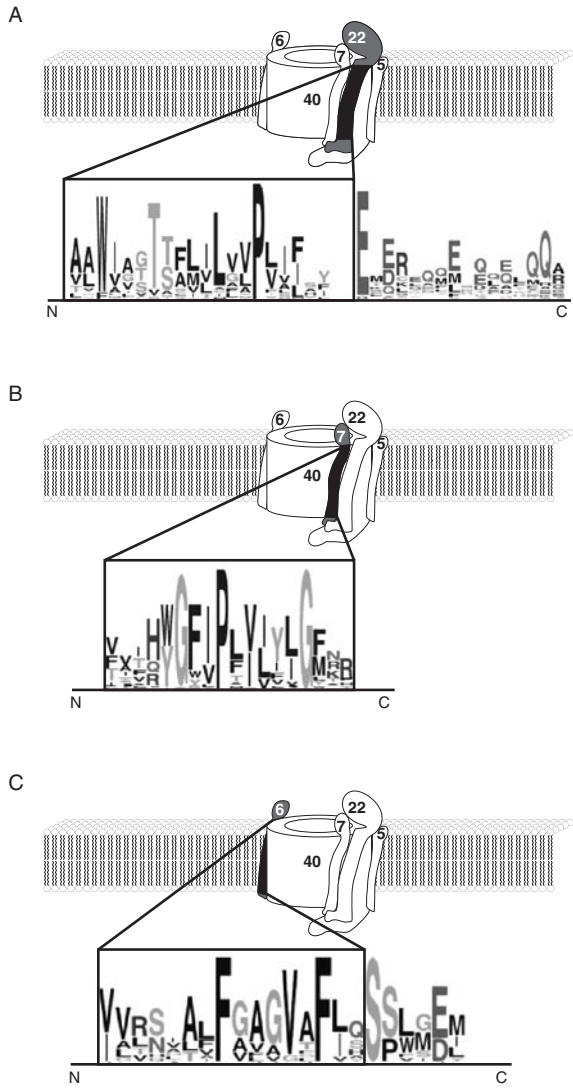



FIG. 12.2. Topology of the highly conserved sequences within the transmembrane segments of (A) Tom22, (B) Tom7, and (C) Tom6. The height of each residue in the WebLogo plot represents its conservation. A conserved proline is present in both the Tom22 and Tom7 transmembrane segments. A similar spacing of aromatic residues is evident in the transmembrane segments of Tom7 and Tom6. Boxes designate the transmembrane segment of the protein.

The intermembrane space domain of Tom22 is rich in negative amino acids, and it was speculated that this region acted to aid precursors to exit the translocation channel [74]. A number of different studies have attempted to further characterize the function of the intermembrane space domain. Some studies of C-terminal truncation mutants lacking the intermembrane space domain showed no effect on the localization of the protein, no diminished growth [87], and no defect observable by *in vitro* protein import [88]. However, another truncation mutant study reported a significant defect [89]. When heterozygous Tom22/Tom22-truncation yeasts are sporulated and dissected, less than 5% of the Tom22-truncation are viable. Of the viable progeny, many rapidly become petite (i.e., lose capability for respiration). When strains are selected that can respire, even these show defects in protein import [89]. This apparent controversy prompted a detailed analysis of the issue by Moczko and colleagues, concluding that a consistent growth defect is observed in three different C-terminal truncation mutants [73]. In the case of each truncation, an ~30% decrease in the import rate of presequence-bearing proteins, but not a carrier protein, supported the hypothesis that Tom22 facilitates the transfer of presequence-bearing proteins through the TOM complex [73].

The importance of the C-terminal domain of Tom22 is further underscored by observations made from comparative genomics. All known Tom22 sequences, from fungi, animals, plants, and from various groups of protists, have a C-terminal domain that would be in the intermembrane space. While there is little conservation of sequence (Figure 12.2A), these domains from all Tom22s are 4–5 kDa in size and rich in amino acids that give them a high propensity for native disorder, including several acidic residues previously suggested to assist the binding of substrates [58, 75]. The intermembrane space domain of Tom22 shares an overlapping function with Tom7 [166] since the coexpression of a Tom22 C-terminal truncation mutant in a Δ*tom7* background shows a severe defect in respiration and is not viable on fermentable carbon sources at 37 °C [166]. Together, Tom22 and Tom7 might contribute to the *trans* site, assisting precursors to exit the channel in Tom40 (Figure 12.3). The Tom22 intermembrane space domain is also the site at which Tim21 binds, forming a TOM–TIM23 complex [90]. This occurs after imported, matrix-destined precursor protein is present at the *trans* site, and is thought to displace the *trans* site-accumulated protein, allowing it to enter the Tim23 channel [90].

### *3. Tom7*

With clear homologues identified in fungi [91], plants [92], animals [93], and a range of single-celled eukaryotes [58], the sequence conservation seen in Tom7 is extraordinary given its small size. A single transmembrane

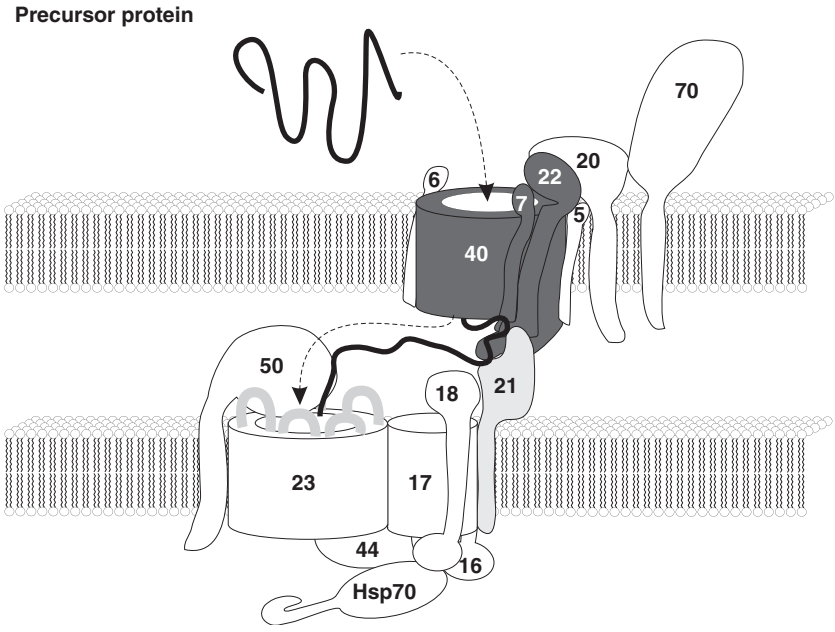



FIG. 12.3. The TOM and TIM23 complexes are transiently tethered via a Tom22–Tim21 interaction. When proteins destined for the TIM23 complex traverse the outer membrane, they interact with the Tom22 intermembrane space domain *trans* site. This *trans* site of the TOM complex is formed from surfaces of Tom40, Tom22, and Tom7. Following the preprotein accumulation at this *trans* site, a Tom22–Tim21 interaction links the two translocation complexes. It is thought that this interaction in the intermembrane space displaces the *trans*-accumulated preprotein, facilitating its entry into the Tim23 channel.

segment is predicted to account for about half the length of these Tom7 proteins [91], with key conserved residues (Figure 12.2B).

Yeast strains from which Tom7 has been deleted are viable on rich growth media, but display slowed growth on nonfermentable carbon sources at higher temperatures and are inviable at 37°C [94]. In *N. crassa*, a knockout of Tom7 shows no apparent growth phenotype [95] but, like the yeast knockout, exhibits defects in the import of some proteins, in particular the outer membrane protein porin [94, 95]. Synthetic phenotypes are observed when the gene encoding Tom7 is deleted in combination with any of the other TOM proteins [96]. The combined loss of Tom7 and Tom20 is lethal in yeast, whereas a combined loss of Tom7 and Tom6 results in slow growth on nonfermentable carbon sources at 30°C [94]. No synthetic phenotype in observed in yeast lacking both Tom7 and Tom70 [94]. In the course of assembling the TOM complex, Tom7 is incorporated into

the 100-kDa assembly intermediate promoting the formation of the mature TOM complex [97, 98]. Recent research has revealed that Tom7 also affects the association of Mdm10 with the SAM complex [99]. While many studies show that Tom7 is important for the overall function of the TOM complex, and its strong conservation in sequence suggests its specific function constrains the structure of Tom7, the precise functional role of the protein remains to be determined.

In Δ*tom7* yeast, more Tom40, Tom20, and Tom22 can be coprecipitated than from wild-type mitochondria; when immunoprecipitation is performed with antibodies against either Tom40, Tom22, or Tom20, the amount of the other two bound proteins increased in mitochondria from the Tom7 knockout [94]. This suggests that the absence of Tom7 results in a more stable interaction between Tom20 and Tom22, and between these molecules and Tom40 [94]. Given Tom7 is tightly bound to Tom40 [48], one interpretation of this data is that Tom20 and Tom7 share a site on the outer surface of the Tom40 barrel, close to Tom22. In the absence of Tom7, Tom20 is able to interact more readily with Tom22 and Tom40 due to reduced steric hindrance. The presence of Tom7 might make this interaction more dynamic. It is of interest in this respect that the residues most highly conserved in the transmembrane segment of Tom7 are rather hydrophilic, as are those in the Tom22 transmembrane segment, and these might interact within the otherwise unfavorable environment of the outer membrane bilayer.

### *4. Tom6*

Another small proteolipid subunit of 6 kDa was first identified as a multicopy suppressor of a temperature-sensitive *tom40* mutant, and shown to associate directly with Tom40 through immunoprecipitation [100]. Since referred to as Tom6, this integral membrane protein [100] is anchored via a C-terminal transmembrane segment [101] (Figure 12.2C). Although a Tom6 subunit is present in all fungi, there are no clear homologues of Tom6 in other eukaryotes; in yeast Tom6 function is not essential. This reflects a common theme in multisubunit machines: later additions to the machines often serve important, but nonessential, functions.

Like Tom7, the precise function of Tom6 is not known, but since the TOM complex is less stable in Δ*tom6* cells, one role played by Tom6 is to stabilize the structure [100]. Overexpression of Tom6 can restore import defects in numerous temperature-sensitive alleles of *tom40* [62, 100]. Furthermore, when Tom6 is deleted, the interactions between Tom40 and the receptors Tom20 and Tom70 are weaker and the receptor binding of precursors is increased [102]. It has therefore been proposed that the function of Tom6 is to allow the receptors to stably dock to the core

TOM complex, thereby promoting them to release precursors for translocation [102]. This model is supported by the fact that Tom6 is important for promoting the association of Tom22 with Tom40 [49], and interacts with Tom22 in a dynamic, preprotein-modulated manner [103].

## B. The Holo-TOM Complex: The Addition of Recently Evolved TOM Receptor Proteins

The TOM holo-complex was first identified from detergent-solubilized extracts of mitochondria from *N. crassa* [46] and can be purified for detailed functional and structural analysis [47, 104]. Under suitable solubilization conditions, Tom20 and Tom70 copurify in approximately stoichiometric amounts with the other subunits of the TOM complex [51]. These receptors are thought to have evolved as add-ons to the primitive translocation machinery as an increasingly large and diverse range of genes were lost to the host nuclear genome [19]. Thus, some subunits can be considered modules of the TOM complex, and we here consider how their presence aids efficiency and enhances the capabilities of the translocase. Not all eukaryotes have the same add-on modules; plants appear to have developed novel solutions to the need for these receptors. It remains possible that some animals, plants, and protists have evolved modules not present in fungi.

### *1. Tom20*

Tom20 is a receptor exposed to the cytosol; the cytosolic domain of Tom20 is hydrophilic and anchored by an N-terminal transmembrane segment [105–107]. This mitochondrial import receptor has been shown to recognize proteins with all manner of mitochondrial targeting sequences. The solution structure of the rat Tom20 cytosolic domain reveals a hydrophobic-binding groove for the targeting segments of substrate proteins [108]. This provides for a recognition step, ahead of the transfer of substrate protein to the Tom40 import channel. Tom20 has a tetratricopeptide repeat (TPR) segment [107, 108]. TPR segments are 34-amino acid structural motifs defined by the consensus sequence $W_4G_8Y_{11}G_{15}Y_{17}A_{20}Y_{24}A_{27}P_{32}$, where each TPR forms a helix-loop-helix structure [109]. This motif is found in over 300 proteins [110] from a diversity of organisms, with tandem arrays of TPRs involved in many cellular processes [111, 112]. The TPR motif is commonly a site of protein–protein interaction and multiprotein assembly [109]. Tom20 remains the only protein known to carry a single TPR segment: in all other know cases, TPRs exist in multiple repeats within a protein [109]. The integrity of the TPR motif is essential for Tom20 to interact with Tom70 [113].

The structure of both a mammalian and plant Tom20 have been elucidated, revealing a unique aspect to the evolution of this receptor. While plants and animals exhibit the presence of a TOM receptor of ~20 kDa with a high affinity for basic amphipathic helices, these two classes of Tom20s have arisen from different ancestral molecules and converged on a very similar solution to the same evolutionary problem. The evidence for this convergent evolution lies in the fact that while the plant and rat Tom20s share similar folds, the plant molecule is anchored to the outer membrane by a C-terminal transmembrane segment whereas the animal Tom20 is N-terminally anchored [53].

Bioinformatics has been used to trace the ancestry of the animal Tom20, showing that the Tom20 found in fungi and in animals are related and presumably derived soon after the split of the last common ancestor to these two groups from other eukaryotes. Many animals have two isoforms of Tom20, with one copy of Tom20 being expressed with only a very limited tissue distribution [114, 115].

Substrate proteins bind coincidently to Tom22 and Tom20 [34, 116, 117], with the mitochondrial presequence able to be cross-linked to these receptors [118]. Tom20 binds a number of precursors with positive N-terminal-targeting sequences, such as Su9-DHFR, but increased salt concentration or mitochondrial presequence peptide can disrupt this binding to the extent that it is comparable to binding of these preproteins to mitochondria lacking Tom20 [117]. However, not all precursor proteins show rate-limiting binding to Tom20: there is a spectrum binding affinities for TOM receptors observed among the various mitochondria proteins, and these receptors cooperate to facilitate import [117, 119]. Similar studies with Tom22 have shown the same salt sensitivity and comparable presequence-binding properties as Tom20. Combined with the results of cross-linking showing Tom20 and Tom22 share the binding of substrates, it seems clear that these receptors act in cooperative manner [116]. Unlike the subsequent binding of substrate proteins to Tom40, it is salt-disruptable electrostatic interactions that promote presequence binding to the Tom20/Tom22 receptor pair; this ionic receptor–presequence interaction is the basis of the acid chain hypothesis [34, 35].

The deletion of Tom20 leads to an initial inability for yeast lacking the gene to survive on nonfermentable carbon sources [107]; this demonstrates that respiration has been seriously compromised in these cells. This defect is more serious than the loss of many other TOM components and, taken with the lower levels of cytochrome *b* and cytochrome *aa3*, Tom20 interactions seem rate-limiting for the import of some respiratory proteins *in vivo* [107]. However, after ~10 days, the Δ*tom20* strain will regain the ability to grow on nonfermentable carbon and mitochondria will import precursors

at wild-type rates. This suppression was shown to affect the mitochondria themselves [120]. Many of the phenotypes initially shown by the Δ*tom20* cells are not due to the loss of Tom20, but a decrease in the steady-state level of Tom22 that results when Tom20 is deleted [74]. The recovery of the Tom20 deletion strain occurs due to a restoration of Tom22 levels in these yeast [74]. It is likely then that the loss of Tom20 causes a destabilization of Tom22, leading to the observed phenotypes. Taken together with the binding assays and cross-linking experiments detailed above, this suggests that Tom22 might represent an ancestral receptor for protein import, with the ''newer'' Tom20 providing a means to improve substrate binding to (and stabilization of) Tom22.

### 2. *Tom70*

A further receptor module of the TOM complex is provided by Tom70. This 70-kDa protein is anchored to the outer mitochondrial membrane by a 29-amino acid N-terminal transmembrane domain [121], which is sufficient to act as a mitochondrial targeting sequence [122, 123]. The Tom70 cytosolic domain contains 11 TPR motifs [52], with the TPRs of the C-terminal domain thought to form a peptide-binding channel [124]. Studies with a 213-residue 25-kDa core domain of Tom70 show that, when just five TPR motifs are present, binding of cleavable and noncleavable preproteins with internal targeting information occurs [125]. The presence of multiple TPRs in Tom70 is understood to provide more than one binding site for preproteins [52, 125].

Many Tom70 substrates are highly hydrophobic and travel to the receptor bound to cytosolic chaperones such as heat shock protein 70 kDa (Hsp70), Hsp90, and 14–3-3 proteins [31, 126]. Recent work on Tom70 has shown that Hsp70 and, at least in humans, Hsp90 bind to a specific site on the first three TPR segments of Tom70 to deliver substrates [31]. The discovery that Tom70 is required for efficient import of the ATP/ADP carrier (AAC) [127, 128] shows that this receptor is important for the import of hydrophobic carrier proteins [52, 128]. In the absence of Tom70, AAC will import at a decreased rate and this is further reduced when antibodies are used to block the Tom20 receptor [128]. The import of a precursor may be rate-limited by a given precursor–receptor interaction due to the chemical properties provided by different receptors, but this does not mean that this rate-limiting interaction is the only important step in substrate recognition. Tom70 is needed for the efficient import of hydrophobic proteins, but these precursors also interact with other receptors such as Tom20 and Tom22.

No Tom70 homologue has been identified in plants or protozoans. This is surprising given that Tom70 is a conserved element in all mammals and

fungi, and many important mitochondrial substrates are recognized by this receptor [129]. The identification of a multiple TPR-containing outer mitochondrial membrane protein of ~64 kDa in *Arabidopsis thaliana* [130] has raised the possibility that this molecule may represent the plant version of the molecule (Figure 12.4).

That evolution continues to finds ways to add modules to molecular machines like the TOM complex can also be seen in the acquisition of Tom71. Identified as a TPR-rich protein in yeast with homology to Tom70 [131, 132], comparative genome analysis shows that Tom71 is a paralogue and the product of a relatively recent gene duplication event in the genomes of the *Saccharomyces*-related clade of yeasts.

### *3. Tom5*

The smallest subunit of the TOM complex in yeast [96], the 50-residue peptide Tom5, is anchored in the outer mitochondrial membrane by an α-helical transmembrane segment located at the C-terminus [96]. This transmembrane segment is crucial for function [133]. Exposed to the cytosol is an α-helical acidic residue-rich N-terminus [96, 134]. Tom5 has been identified in a range of species of fungi, but does not appear to be present in the TOM complex of plants, animals, or other eukaryotes.

The *TOM5* gene is not essential for yeast growth at 30°C, with the knockout displaying a reduced growth rate on both fermentable and non-fermentable carbon when compared to a wild-type strain [96]. Mitochondria isolated from Δ*tom5* yeast show severe defects in the import of mitochondrial proteins which reside in all four mitochondrial compartments [96]. However, these findings do not appear to apply more generally, as the loss of Tom5 in *N. crassa* does not exhibit the import and growth defects observed in *S. cerevisiae* [50].

The specific function of Tom5 has been difficult to define. Although early reports have shown that the deletion of Tom5 alone in yeast does not appear to reveal a destabilization of the TOM complex [49, 96], a recent study has claimed that the yeast TOM complex is less stable in the absence of Tom5 [50]. Even in *N. crassa*, it appears that a loss of Tom5 may have a destabilizing effect on the TOM complex, with the study of a Δ*tom5*, Δ*tom6* deletion in *N. crassa* implicating a role for Tom5 in TOM complex stabilization: the TOM complex in a Δ*tom5*,Δ*tom6* strain is less stable than in the Δ*tom6* strain [95]. These findings are significant given the role that Tom5 plays in the biogenesis of the TOM complex. When the assembly of newly imported Tom40 subunits is tracked using blue-native polyacrylamide gel electrophoresis, Tom5 is first observed in a 100-kDa assembly intermediate complex with Tom40 [98], as well as the final 400-kDa mature

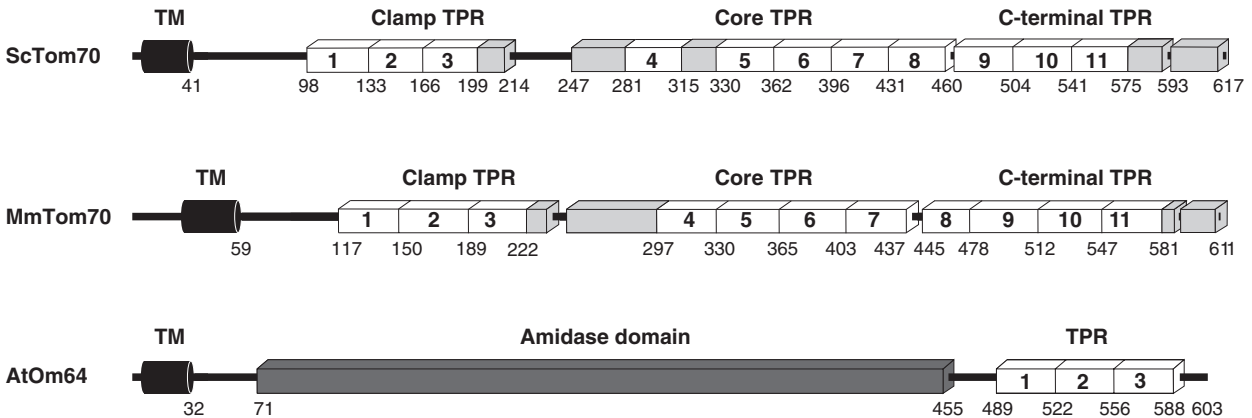



FIG. 12.4. The domain structure of Tom70. Conservation of the domain organization with 11 TPR segments of Tom70 is evident in the yeast (ScTom70) and mouse (MmTom70) proteins. The *A. thaliana* protein OM64 (AtOm64), with three C-terminal TPR segments, represents the best candidate to function as an equivalent import receptor for the plant TOM complex.

TOM complex [97]. This 100-kDa complex is the first recognizable form taken on by the newly folded Tom40 subunits as they leave the SAM complex [97, 98]. Thus, Tom5 might represent a recent module added to the TOM complex in fungi to assist the recruitment of other TOM subunits onto newly assembled Tom40 barrels.

## IV. The SAM Complex

### A. Machinery for the Assembly of Complex Proteins into the Mitochondrial Outer Membrane

The majority of proteins in the mitochondrial outer membrane appear to be of relatively simple topology, with a single transmembrane anchor region critical for sorting and insertion [43–45]. Some of these proteins have their transmembrane segment at the N-terminus and are referred to as signal-anchor proteins [29]. The model for signal-anchor protein insertion is one in which the molecules are recognized by the TOM complex, but do not require access to the Tom40 channel or the presence of TOM receptors [64, 106, 122, 135]. This suggests that the core TOM complex recognizes these proteins and inserts them into the membrane. It is likely that tail-anchored proteins, with a single transmembrane segment at their C-terminus, are imported in the same way. The mechanism behind these protein insertion pathways is not yet known.

Mitochondria were derived from α-proteobacteria and several key outer membrane proteins retain a bacterial β-barrel topology. How β-barrel proteins, and in particular the TOM complex itself, assemble in the outer membrane had remained perplexing for some time. In recent years, a second mitochondrial outer membrane complex involved in protein import has been identified [98]. This SAM complex was discovered through investigations into the enigmatic protein Mas37, which was renamed Sam37 when it was shown to function in the SAM complex. Since the β-barrel assembly process is essential and Sam37 is not an essential protein, it seemed likely that other subunits of the SAM complex remained to be discovered.

Homologues to the bacterial protein Omp85 were identified and shown to have an essential role in mitochondrial outer membrane biogenesis [136–139]. This β-barrel protein, Sam50, forms the core of the SAM complex. Peripherally associated with Sam50 on the cytosolic side of the membrane are the associated proteins Sam35 and Sam37. Mdm10 also participates with this complex, perhaps as a more transiently associated module.

The SAM complex is not an import channel, but rather acts downstream of TOM complex-mediated import to integrate proteins of complex

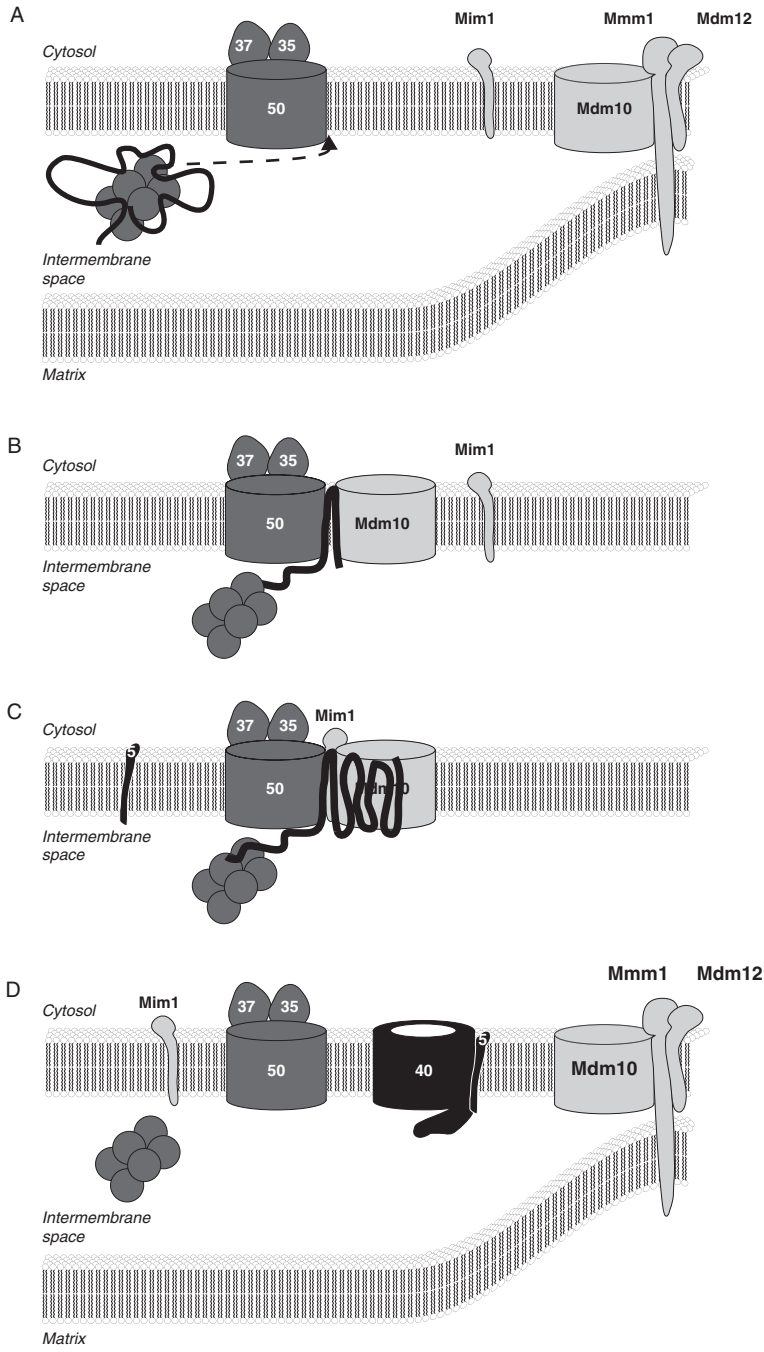
topology into the mitochondrial outer membrane. The way in which the SAM complex acts to insert and assemble β-barrel proteins has not yet been elucidated, but two models have been put forward. It was speculated that a pore within Sam50 may be the site at which β-barrels were assembled prior to a lateral release event in which Sam50 would unfold and open to release the precursor into the outer mitochondrial membrane. However, there would be unfavorable thermodynamic consequences incurred in the unfolding of a highly stable β-barrel protein like Sam50 [65]. An alternative model proposes that after passing through the TOM complex, precursors bind the small Tim chaperones in the intermembrane space prior to delivery to the SAM complex. Here it is thought that the SAM complex acts as a scaffold to disrupt the lipid bilayer and allow the protein to enter the membrane [136, 137] (Figure 12.5). In the case of TOM complex assembly, increasing evidence is mounting to suggest that other proteins (such as Mim1) are critical to this process. How the SAM complex assists the assembly of the TOM complex, and how it may interact with other ''TOM assembly'' molecules, is a hot topic of research.

### *1. Sam50*

The widespread conservation seen between Sam50 and the Gram-negative bacterial protein Omp85 [136, 138, 139] sparked great interest in its function. Early clues were provided in the form of two publications on the bacterial Omp85: one showed the molecule to be important for the integration of (β-barrel) proteins in the outer membrane of Gram-negative bacteria [140], and the other claimed a role for Omp85 in membrane lipid biosynthesis [141].

The mitochondrial Sam50 is an ~50-kDa β-barrel with a conserved domain structure: a β-barrel is predicted at the C-terminus, and the N-terminus contains one predicted POlypeptide-TRAnsport-associated (POTRA) domain. The POTRA domain is found in a number of β-barrel proteins involved in protein translocation [137, 142]. While the function of this domain has not been fully elucidated, it is speculated to act in a chaperone-like manner to assist translocation [142]. Depletion of Sam50 in yeast results in the defective import of Tom40 and VDAC, but does not impair the import of single-transmembrane spanning proteins [136] and matrix proteins [138].

The import of Sam50 into mitochondria has been studied. Nascent molecules of Sam50 are in an unfolded state when recognized at the TOM complex. This recognition requires the receptors Tom20 and Tom70; absence of either will result in a reduction in imported Sam50 [143]. After translocation through the Tom40 channel, unfolded Sam50 interacts with the small Tim chaperones (Tim9/10 and Tim8/13) of the intermembrane


A
Mim1
Mmm1
Mdm12
Cytosol
37
35
50
Mdm10
Intermembrane space
Matrix
B
Mim1
Cytosol
37
35
50
Mdm10
Intermembrane space
C
Mim1
Cytosol
37
35
50
Intermembrane space
D
Mmm1
Mdm12
Mim1
Cytosol
37
35
50
40
Mdm10
Intermembrane space
Matrix

space [143]. The unfolded Sam50 is then delivered to a preexisting SAM complex for integration. The role of existing SAM complexes is supported by findings that mitochondria depleted of Sam50 or Sam35 show significant defects in Sam50 insertion [143]. In contrast, the absence of Sam37 does not result in as severe a decrease in Sam50 assembly [143].

### 2. *Sam37*

The Sam37 protein was first identified as the result of a yeast mutant screen for temperature-sensitive growth on nonfermentable carbon sources [144]. This research aimed to identify genes relating to phospholipid biosynthesis, with a loss of Sam37 resulting in a reduction in the amounts of mitochondria-synthesized lipids [144]. Early observation of protein import defects in this gene knockout were similar to those observed in Δ*tom70* yeast: some proteins are imported at wild-type rates, whereas AAC import is deficient [52, 144]. This led to the suggestion that Sam37 may be another receptor subunit of the TOM complex. However, the yeast Sam37 does not migrate with the TOM complex on blue native PAGE and is not required for binding AAC, but affects AAC import indirectly of this event [145]. Subsequently, Sam37 was inferred to be peripherally associated with Sam50 on the outer mitochondrial membrane [98, 145], based on its extractability with sodium carbonate. This ''alkali extraction'' assay is not reliable when applied to mitochondrial outer membrane proteins [43], though it is clear from protease protection assays that the bulk of Sam37 is exposed on the cytosolic face of the SAM complex [98, 144]. Perhaps because of this extrinsic topology, Sam37 is not imported through the TOM complex but might instead assemble directly onto existing Sam50 subunits [143].

There remains some controversy over the existence of a mammalian Sam37. The best candidate is Metaxin-1, an ∼35-kDa protein localized to the outer mitochondrial membrane by a sequence at its C-terminus [146]. Metaxin-1 shows similarity to Sam37 only in the N-terminal region, where there is a 25% sequence identity [146]. This low level of sequence similarity makes unsurprising its failure to complement for the loss of *sam37* in

FIG. 12.5. A model for the insertion and assembly of Tom40 by the SAM complex. After translocation by the TOM complex (not shown), imported Tom40 is transferred to the SAM complex via the small Tim proteins of the intermembrane space (A). Unfolded Tom40 is inserted into the outer membrane, a process likely to be mediated by the transmembrane surfaces of the SAM complex (B). Tom40 assembles within the membrane, in complex with SAM components (C); Mim1 appears to act at a late stage to assist Tom40 to assemble with other TOM components. Tom40 releases from the SAM complex, forming the second assembly intermediate; this requires the addition of Tom5 to the Tom40 barrel (D).

yeast [146]. Like Sam37, Metaxin-1 appears to contribute to protein import into mitochondria: antibodies against Metaxin-1 will severely reduce the import of the mitochondrial protein preadrenodoxin [146].

From early observations, Tom70 and Sam37 appeared to form a heterodimer [144], and interact genetically: overexpression of either Tom70 or Sam37 is enhanced only if the other is also upregulated [144]. It is therefore tempting to speculate that the SAM and TOM complexes can form a ''supercomplex'' in the outer mitochondrial membrane, with interactions mediated by these two components. It remains to be seen whether this supercomplex exists.

### *3. Sam35*

A second ''peripherally'' associated subunit of the SAM complex, Sam35, is an essential protein [147, 148]. Like Sam37, Sam35 shows some sequence similarity to a member of the metaxin family, in this case Metaxin-2. In animals, Metaxin-1 and Metaxin-2 interact with each other, though it is not yet clear that interactions are made between the metaxins and the mammalian Sam50 [149].

A temperature-sensitive mutant of Sam35 was used to characterize the function of this protein. When *in vitro* translated Tom40 is imported into *sam35* mutant mitochondria, the assembly of Tom40 is retarded [147]. Rather than the formation of the second assembly intermediate (100 kDa) and mature TOM complex (400 kDa), Tom40 remains stuck at the first assembly intermediate [147]. This assembly intermediate is known to contain Sam37 and Sam50: it is the SAM complex-Tom40 intermediate formed during Tom40 import and assembly [147]. Furthermore, the effect of the temperature-sensitive Sam35 mutation is only observed for the import of β-barrel proteins Tom40 and porin; the import of presequence-bearing proteins and carrier proteins is not affected [147]. As with the temperature-sensitive Sam35 mutant, shutting down the expression of Sam35 leads to a block in the biogenesis of functional TOM complexes [148].

## B. Additional Modules of the SAM in the Outer Membrane

### *1. Mdm10*

First identified as a protein with a role in the maintenance of mitochondrial distribution and morphology [150], Mdm10 is part of a complex with Mdm12 and Mmm1, which functions to attach yeast mitochondria to the actin cytoskeleton [151]. None of these fungal proteins appear to have obvious homologues in animals or other eukaryotes (Lithgow, unpublished data). It was shown that Mdm10 also plays an important role in

assisting the SAM complex for protein assembly into the outer membrane (Figure 12.5) [152]. Conceptually, Mdm10 (and perhaps Mmm1 and Mdm12) might represent a distinct module of the SAM complex.

Mdm10 is an ~56-kDa protein [150] of the outer mitochondrial membrane with a predicted β-barrel structure. A tagged version of Sam37 was used to demonstrate that Mdm10 can be copurified with the SAM complex by affinity chromatography [152]. The presence of Mdm10 is important in the assembly of Tom40, with studies using Δ*mdm10* yeast revealing a defect in the assembly of Tom40, but not porin [152]. Neither does a loss of Mdm10 alter import rates for proteins destined for the inner membrane or matrix compartments [152]. Thus, while Mdm10 assists Tom40 to assemble with Tom22, Tom6, and Tom7 (subsequent to its association with Tom5) [152], it remains to be determined precisely what role Mdm10 plays in this process.

### 2. *Mim1*

An integral outer membrane protein, termed Mim1, has been identified [148, 153, 154] and is found in a number of yeast species. Mim1 is not essential, but its deletion causes a severe respiratory deficiency in yeast. Under the standard conditions applied for solubilizing the SAM complex from the mitochondrial outer membrane, Mim1 does not copurify with the SAM complex by immunoprecipitation, gel filtration, or density gradient centrifugation [148]. However, the function of Mim1 is to assist assembly of the TOM complex [148, 153, 154], suggesting it might be a transiently associating module of the SAM complex. The depletion of Mim1 results in a Tom40 assembly defect similar to that observed for the deletion of Mdm10, with blue-native PAGE analysis showing that the *in vitro* assembly of the mature TOM complex is inhibited in Mim1-depleted mitochondria [148, 152, 153]. FLAG-tagged Mim1 appears to exist in a 180-kDa complex [148]; whether this represents a module of the SAM complex remains to be seen. It is of interest that this Mim1 complex is approximately the size at which Mdm10 runs when *not* in the SAM holocomplex [99].

While Tom40 assembly is assisted by Mim1, depletion of Mim1 does not appear to affect the *in vitro* import and assembly of porin [148, 153]. Due to the conservation of residues in the predicted transmembrane segment of Mim1, it has been proposed that this protein may specifically assist Tom40 to correctly assemble with the TOM receptors [153]. Only very few sequences of Mim1 proteins are available, but the conserved residues do share some similarity to those conserved in the transmembrane segments of Tom22 and Tom7, which might provide for specific, transient contacts between Mim1 and Tom40, and any subunit with a greater affinity (e.g., Tom22) might subsequently displace the Mim1 ''chaperone.''

## V. Concluding Remarks

The molecular machines for protein import across the outer mitochondrial membrane demonstrate a modular structure. The apparent paradox of Mdm10, a morphology protein with a role in TOM complex biogenesis, is in fact but one of many examples in which a module of proteins can be recruited to a complex to add-on a function. The presence of Mdm10 as a component of both the SAM complex and the Mdm10/12/Mmm1 complex can be considered in the context of the work of Gavin and colleagues, who have presented a new paradigm for the organization of eukaryotic molecular machines. Their research shows that multisubunit protein complexes are common and that protein ''modules''(defined as two or more proteins found together in multiple complexes) can be recruited to a number of different complexes in order to perform a specialized function [155]; this system allows complexes to perform multiple functions. An example from this survey of yeast protein machinery is that the protein lipoamide dehydrogenase is present in two machines of similar phenotype (the pyruvate complex and the α-ketogluterate complex) [155]. This is but one example in which a shared protein can play a wider role than the specific complexes in which it is present [155].

The modularity of protein complexes described by Gavin *et al.* [155] is also seen in the protein import machinery of the mitochondrial outer membrane. The TOM complex consists of a stable, highly conserved core to which receptor modules can dock to permit efficient precursor recognition and import. Moreover, the dual role of Mdm10 is an example of a protein functioning in two different contexts. It would be interesting to consider whether the TOM and SAM complexes also act as modules of a transient ''supercomplex.''

The origins of the TOM and SAM import proteins and their evolution are aspects of these machines which are still largely unexplored. Mitochondria arose from bacteria, and undoubtedly members of the mitochondrial protein import machines evolved from bacterial proteins. This is true of Sam50: new protein modules evolved to convert the protomitochondiral Omp85 into a highly efficient complex for β-barrel assembly. While no bacterial protein with strong sequence similarity to Tom40 has been found, the β-barrel structure infers a bacterial origin. It is possible that the evolution of this protein has become so divergent as to make its precursor form unrecognizable in present-day bacteria.

Just as mitochondria are bacteria-derived organelles, and some of their protein translocation machineries are bacteria derived, they can be the target of attack by pathogenic bacteria. A number of bacterial and viral

pathogens produce proteins that are targeted to mitochondria as part of the disease-causing process [156]. These include the Map and EspF proteins of enteropathogenic *E. coli*, which are imported via the TOM complex [157–159]. Map has an N-terminal targeting sequence and is targeted to the matrix, affecting mitochondrial morphology [158]. In contrast, the protein EspF disrupts the transmembrane potential across the inner membrane, and triggers cell death [157, 159]. The porin PorB from *Neisseria gonorrhoeae* and *Neisseria meningitidis* is also targeted to the mitochondria of infected cells to affect apoptosis [160–162]. Further examples include the apoptosis-inducing protein Omp38 from *Acinetobacter baumannii* [163], the VacA protein of *Helicobacter pylori* [164], and proteins from a host of viruses including *Hepatitis B virus, Human papillomavirus,* and human T-cell leukaemia virus type 1 [165]. These recent examples show that the molecular machinery of the outer mitochondrial membrane is not only essential for the survival of the host cell but, by virtue of its exploitation as the entry point for bacteria-derived proteins, it is also critical to the success of a diverse array of human pathogens.

## Acknowledgments

Our work is supported by an Australian Postgraduate Award (to MJD) and a grant from the Australian Research Council (to TL). We thank members of the Lithgow laboratory for critical comments on the chapter.

## References

1. Schatz, G. (1995). Mitochondria: beyond oxidative phosphorylation. *Biochim Biophys Acta* 1271:123–126.
2. Bartlett, K., and Eaton, S. (2004). Mitochondrial β-oxidation. *Eur J Biochem* 271:462–469.
3. Kohlhaw, G.B. (2003). Leucine biosynthesis in fungi: entering metabolism through the back door. *Microbiol Mol Biol Rev* 67:1–15.
4. Lill, R., and Muhlenhoff, U. (2005). Iron-sulfur-protein biogenesis in eukaryotes. *Trends Biochem Sci* 30:133–141.
5. Ekert, P.G., and Vaux, D.L. (2005). The mitochondrial death squad: hardened killers or innocent bystanders? *Curr Opin Cell Biol* 17:626–630.
6. Armstrong, J.S. (2006). Mitochondrial membrane permeabilization: the sine qua non for cell death. *Bioessays* 28:253–260.
7. Lardy, H.A., and Ferguson, S.M. (1969). Oxidative phosphorylation in mitochondria. *Annu Rev Biochem* 38:991–1034.
8. Saraste, M. (1999). Oxidative phosphorylation at the fin de siecle. *Science* 283:1488–1493.
9. Martin, W., and Muller, M. (1998). The hydrogen hypothesis for the first eukaryote. *Nature* 392:37–41.

10. Dolezal, P., and Lithgow, T. (2006). Evolution of the molecular machines for protein import into mitochondria. *Science* 313:314–318.
11. Meisinger, C., Brix, J., Model, K., Pfanner, N., and Ryan, M.T. (1999). The preprotein translocase of the outer mitochondrial membrane: receptors and a general import pore. *Cell Mol Life Sci* 56:817–824.
12. Pfanner, N., and Chacinska, A. (2002). The mitochondrial import machinery: preprotein-conducting channels with binding sites for presequences. *Biochim Biophys Acta* 1592:15–24.
13. Taylor, R.D., and Pfanner, N. (2004). The protein import and assembly machinery of the mitochondrial outer membrane. *Biochim Biophys Acta* 1658:37–43.
14. Pfanner, N., Wiedemann, N., Meisinger, C., and Lithgow, T. (2004). Assembling the mitochondrial outer membrane. *Nat Struct Mol Biol* 11:1044–1048.
15. Paschen, S.A., Neupert, W., and Rapaport, D. (2005). Biogenesis of beta-barrel membrane proteins of mitochondria. *Trends Biochem Sci* 30:575–582.
16. Rehling, P., Brandner, K., and Pfanner, N. (2004). Mitochondrial import and the twin-pore translocase. *Nat Rev Mol Cell Biol* 5:519–530.
17. Koehler, C.M. (2004). New developments in mitochondrial assembly. *Annu Rev Cell Dev Biol* 20:309–335.
18. Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J.M. (2005). A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121:1059–1069.
19. Lucattini, R., Likic, V.A., and Lithgow, T. (2004). Bacterial proteins predisposed for targeting to mitochondria. *Mol Biol Evol* 21:652–658.
20. Baker, A., and Schatz, G. (1987). Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. *Proc Natl Acad Sci USA* 84:3117–3121.
21. Guda, C., Guda, P., Fahy, E., and Subramaniam, S. (2004). MITOPRED: a web server for the prediction of mitochondrial proteins. *Nucleic Acids Res* 32:W372–W374.
22. Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4:1581–1590.
23. Claros, M.G., Perea, J., Shu, Y., Samatey, F.A., Popot, J.L., and Jacq, C. (1995). Limitations to *in vivo* import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome b. *Eur J Biochem* 228:762–771.
24. Lain, B., Iriarte, A., Mattingly, J.R., Jr., Moreno, J.I., and Martinez-Carrion, M. (1995). Structural features of the precursor to mitochondrial aspartate aminotransferase responsible for binding to hsp70. *J Biol Chem* 270:24732–24739.
25. Lain, B., Iriarte, A., and Martinez-Carrion, M. (1994). Dependence of the folding and import of the precursor to mitochondrial aspartate aminotransferase on the nature of the cell-free translation system. *J Biol Chem* 269:15588–15596.
26. Cyr, D.M., and Neupert, W. (1996). Roles for hsp70 in protein translocation across membranes of organelles. *EXS* 77:25–40.
27. Stuart, R.A., Cyr, D.M., Craig, E.A., and Neupert, W. (1994). Mitochondrial molecular chaperones: their role in protein translocation. *Trends Biochem Sci* 19:87–92.
28. Murakami, H., Pain, D., and Blobel, G. (1988). 70-kD heat shock-related protein is one of at least two distinct cytosolic factors stimulating protein import into mitochondria. *J Cell Biol* 107:2051–2057.
29. Shore, G.C., McBride, H.M., Millar, D.G., Steenaart, N.A., and Nguyen, M. (1995). Import and insertion of proteins into the mitochondrial outer membrane. *Eur J Biochem* 227:9–18.

30. Hachiya, N., Komiya, T., Alam, R., Iwahashi, J., Sakaguchi, M., Omura, T., and Mihara, K. (1994). MSF, a novel cytoplasmic chaperone which functions in precursor targeting to mitochondria. *EMBO J* 13:5146–5154.
31. Young, J.C., Hoogenraad, N.J., and Hartl, F.U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell* 112:41–50.
32. Beddoe, T., and Lithgow, T. (2002). Delivery of nascent polypeptides to the mitochondrial surface. *Biochim Biophys Acta* 1592:35–39.
33. Voos, W., and Rottgers, K. (2002). Molecular chaperones as essential mediators of mitochondrial biogenesis. *Biochim Biophys Acta* 1592:51–62.
34. Honlinger, A., Kubrich, M., Moczko, M., Gartner, F., Mallet, L., Bussereau, F., Eckerskorn, C., Lottspeich, F., Dietmeier, K., Jacquet, M., and Pfanner, N. (1995). The mitochondrial receptor complex: Mom22 is essential for cell viability and directly interacts with preproteins. *Mol Cell Biol* 15:3382–3389.
35. Schatz, G. (1997). Just follow the acid chain. *Nature* 388:121–122.
36. Rapaport, D. (2003). Finding the right organelle. Targeting signals in mitochondrial outer-membrane proteins. *EMBO Rep* 4:948–952.
37. Nakai, M., Takeda, A., Cleary, M.L., and Endo, T. (1993). The bcl-2 protein is inserted into the outer membrane but not into the inner membrane of rat liver mitochondria *in vitro*. *Biochem Biophys Res Commun* 196:233–239.
38. Krajewski, S., Tanaka, S., Takayama, S., Schibler, M.J., Fenton, W., and Reed, J.C. (1993). Investigation of the subcellular distribution of the bcl-2 oncoprotein: residence in the nuclear envelope, endoplasmic reticulum, and outer mitochondrial membranes. *Cancer Res* 53:4701–4714.
39. Beilharz, T., Egan, B., Silver, P.A., Hofmann, K., and Lithgow, T. (2003). Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *J Biol Chem* 278:8219–8223.
40. Kanaji, S., Iwahashi, J., Kida, Y., Sakaguchi, M., and Mihara, K. (2000). Characterization of the signal that directs Tom20 to the mitochondrial outer membrane. *J Cell Biol* 151:277–288.
41. Isenmann, S., Khew-Goodall, Y., Gamble, J., Vadas, M., and Wattenberg, B.W. (1998). A splice-isoform of vesicle-associated membrane protein-1 (VAMP-1) contains a mitochondrial targeting signal. *Mol Biol Cell* 9:1649–1660.
42. Sirrenberg, C., Endres, M., Becker, K., Bauer, M.F., Walther, E., Neupert, W., and Brunner, M. (1997). Functional cooperation and stoichiometry of protein translocases of the outer and inner membranes of mitochondria. *J Biol Chem* 272:29963–29966.
43. Burri, L., Vascotto, K., Gentle, I.E., Chan, N.C., Beilharz, T., Stapleton, D.I., Ramage, L., and Lithgow, T. (2006). Integral membrane proteins in the mitochondrial outer membrane of *Saccharomyces cerevisiae*. *FEBS J* 273:1507–1515.
44. Schmitt, S., Prokisch, H., Schlunck, T., Camp, D.G., II, Ahting, U., Waizenegger, T., Scharfe, C., Meitinger, T., Imhof, A., Neupert, W., Oefner, P.J., and Rapaport, D. (2006). Proteome analysis of mitochondrial outer membrane from *Neurospora crassa*. *Proteomics* 6:72–80.
45. Zahedi, R.P., Sickmann, A., Boehm, A.M., Winkler, C., Zufall, N., Schonfisch, B., Guiard, B., Pfanner, N., and Meisinger, C. (2006). Proteomic analysis of the yeast mitochondrial outer membrane reveals accumulation of a subclass of preproteins. *Mol Biol Cell* 17:1436–1450.
46. Kiebler, M., Pfaller, R., Sollner, T., Griffiths, G., Horstmann, H., Pfanner, N., and Neupert, W. (1990). Identification of a mitochondrial receptor complex required for recognition and membrane insertion of precursor proteins. *Nature* 348:610–616.

47. Ahting, U., Thun, C., Hegerl, R., Typke, D., Nargang, F.E., Neupert, W., and Nussberger, S. (1999). The TOM core complex: the general protein import pore of the outer membrane of mitochondria. *J Cell Biol* 147:959–968.
48. Meisinger, C., Ryan, M.T., Hill, K., Model, K., Lim, J.H., Sickmann, A., Muller, H., Meyer, H.E., Wagner, R., and Pfanner, N. (2001). Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors. *Mol Cell Biol* 21:2337–2348.
49. Dekker, P.J., Ryan, M.T., Brix, J., Muller, H., Honlinger, A., and Pfanner, N. (1998). Preprotein translocase of the outer mitochondrial membrane: molecular dissection and assembly of the general import pore complex. *Mol Cell Biol* 18:6515–6524.
50. Schmitt, S., Ahting, U., Eichacker, L., Granvogl, B., Go, N.E., Nargang, F.E., Neupert, W., and Nussberger, S. (2005). Role of Tom5 in maintaining the structural stability of the TOM complex of mitochondria. *J Biol Chem* 280:14499–14506.
51. Kunkele, K.P., Heins, S., Dembowski, M., Nargang, F.E., Benz, R., Thieffry, M., Walz, J., Lill, R., Nussberger, S., and Neupert, W. (1998). The preprotein translocation channel of the outer membrane of mitochondria. *Cell* 93:1009–1019.
52. Chan, N.C., Likic, V.A., Waller, R.F., Mulhern, T.D., and Lithgow, T. (2006). The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. *J Mol Biol* 358:1010–1022.
53. Perry, A.J., Hulett, J.M., Likic, V.A., Lithgow, T., and Gooley, P.R. (2006). Convergent evolution of receptors for protein import into mitochondria. *Curr Biol* 16:221–229.
54. Schwartz, M.P., Huang, S., and Matouschek, A. (1999). The structure of precursor proteins during import into mitochondria. *J Biol Chem* 274:12759–12764.
55. Vestweber, D., Brunner, J., Baker, A., and Schatz, G. (1989). A 42K outer-membrane protein is a component of the yeast mitochondrial protein import site. *Nature* 341:205–209.
56. Baker, K.P., Schaniel, A., Vestweber, D., and Schatz, G. (1990). A yeast mitochondrial outer membrane protein essential for protein import and cell viability. *Nature* 348:605–609.
57. Mannella, C.A., Neuwald, A.F., and Lawrence, C.E. (1996). Detection of likely transmembrane beta strand regions in sequences of mitochondrial pore proteins using the Gibbs sampler. *J Bioenerg Biomembr* 28:163–169.
58. Macasev, D., Whelan, J., Newbigin, E., Silva-Filho, M.C., Mulhern, T.D., and Lithgow, T. (2004). Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. *Mol Biol Evol* 21:1557–1564.
59. Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins [see comment]. *Nature* 395:516–521.
60. Schwartz, M.P., and Matouschek, A. (1999). The dimensions of the protein import channels in the outer and inner mitochondrial membranes. *Proc Natl Acad Sci USA* 96:13086–13090.
61. Kunkele, K.P., Juin, P., Pompa, C., Nargang, F.E., Henry, J.P., Neupert, W., Lill, R., and Thieffry, M. (1998). The isolated complex of the translocase of the outer membrane of mitochondria. Characterization of the cation-selective and voltage-gated preprotein-conducting pore. *J Biol Chem* 273:31032–31039.
62. Gabriel, K., Egan, B., and Lithgow, T. (2003). Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins. *EMBO J* 22:2380–2386.

63. Esaki, M., Kanamori, T., Nishikawa, S., Shin, I., Schultz, P.G., and Endo, T. (2003). Tom40 protein import channel binds to non-native proteins and prevents their aggregation. *Nat Struct Biol* 10:988–994.
64. Ahting, U., Waizenegger, T., Neupert, W., and Rapaport, D. (2005). Signal-anchored proteins follow a unique insertion pathway into the outer membrane of mitochondria. *J Biol Chem* 280:48–53.
65. Gabriel, K., Buchanan, S.K., and Lithgow, T. (2001). The alpha and the beta: protein translocation across mitochondrial and plastid outer membranes. *Trends Biochem Sci* 26:36–40.
66. Becker, L., Bannwarth, M., Meisinger, C., Hill, K., Model, K., Krimmer, T., Casadio, R., Truscott, K.N., Schulz, G.E., Pfanner, N., and Wagner, R. (2005). Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore. *J Mol Biol* 353:1011–1020.
67. Rapaport, D., Kunkele, K.P., Dembowski, M., Ahting, U., Nargang, F.E., Neupert, W., and Lill, R. (1998). Dynamics of the TOM complex of mitochondria during binding and translocation of preproteins. *Mol Cell Biol* 18:5256–5262.
68. Model, K., Prinz, T., Ruiz, T., Radermacher, M., Krimmer, T., Kuhlbrandt, W., Pfanner, N., and Meisinger, C. (2002). Protein translocase of the outer mitochondrial membrane: role of import receptors in the structural organization of the TOM complex. *J Mol Biol* 316:657–666.
69. Bains, G., and Lithgow, T. (1999). The Tom channel in the mitochondrial outer membrane: alive and kicking. *Bioessays* 21:1–4.
70. Suzuki, H., Kadowaki, T., Maeda, M., Sasaki, H., Nabekura, J., Sakaguchi, M., and Mihara, K. (2004). Membrane-embedded C-terminal segment of rat mitochondrial TOM40 constitutes protein-conducting pore with enriched beta-structure. *J Biol Chem* 279:50619–50629.
71. Chaudhuri, M., and Nargang, F.E. (2003). Import and assembly of *Neurospora crassa* Tom40 into mitochondria of *Trypanosoma brucei in vivo*. *Curr Genet* 44:85–94.
72. Kiebler, M., Keil, P., Schneider, H., van der Klei, I.J., Pfanner, N., and Neupert, W. (1993). The mitochondrial receptor complex: a central role of MOM22 in mediating preprotein transfer from receptors to the general insertion pore. *Cell* 74:483–492.
73. Moczko, M., Bomer, U., Kubrich, M., Zufall, N., Honlinger, A., and Pfanner, N. (1997). The intermembrane space domain of mitochondrial Tom22 functions as a trans binding site for preproteins with N-terminal targeting sequences. *Mol Cell Biol* 17:6574–6584.
74. Lithgow, T., Junne, T., Suda, K., Gratzer, S., and Schatz, G. (1994). The mitochondrial outer membrane protein Mas22p is essential for protein import and viability of yeast. *Proc Natl Acad Sci USA* 91:11973–11977.
75. Perry, A.J., and Lithgow, T. (2005). Protein targeting: entropy, energetics and modular machines. *Curr Biol* 15:R423–R425.
76. Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem Sci* 27:527–533.
77. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21:3433–3434.
78. Nargang, F.E., Rapaport, D., Ritzel, R.G., Neupert, W., and Lill, R. (1998). Role of the negative charges in the cytosolic domain of TOM22 in the import of precursor proteins into mitochondria. *Mol Cell Biol* 18:3173–3181.
79. Braun, H.P., and Schmitz, U.K. (1999). The protein-import apparatus of plant mitochondria. *Planta* 209:267–274.
80. Lister, R., Hulett, J.M., Lithgow, T., and Whelan, J. (2005). Protein import into mitochondria: origins and functions today (review). *Mol Membr Biol* 22:87–100.

81. Egan, B., Beilharz, T., George, R., Isenmann, S., Gratzer, S., Wattenberg, B., and Lithgow, T. (1999). Targeting of tail-anchored proteins to yeast mitochondria *in vivo*. *FEBS Lett* 451:243–248.
82. Saeki, K., Suzuki, H., Tsuneoka, M., Maeda, M., Iwamoto, R., Hasuwa, H., Shida, S., Takahashi, T., Sakaguchi, M., Endo, T., Miura, Y., Mekada, E., and Mihara, K. (2000). Identification of mammalian TOM22 as a subunit of the preprotein translocase of the mitochondrial outer membrane. *J Biol Chem* 275:31996–32002.
83. Yano, M., Hoogenraad, N., Terada, K., and Mori, M. (2000). Identification and functional analysis of human Tom22 for protein import into mitochondria. *Mol Cell Biol* 20:7205–7213.
84. Ulmschneider, M.B., and Sansom, M.S. (2001). Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta* 1512:1–14.
85. Adamian, L., Nanda, V., DeGrado, W.F., and Liang, J. (2005). Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. *Proteins* 59:496–509.
86. van Wilpe, S., Ryan, M.T., Hill, K., Maarse, A.C., Meisinger, C., Brix, J., Dekker, P.J., Moczko, M., Wagner, R., Meijer, M., Guiard, B., Honlinger, B., *et al.* (1999). Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. *Nature* 401:485–489.
87. Court, D.A., Nargang, F.E., Steiner, H., Hodges, R.S., Neupert, W., and Lill, R. (1996). Role of the intermembrane-space domain of the preprotein receptor Tom22 in protein import into mitochondria. *Mol Cell Biol* 16:4035–4042.
88. Nakai, M., Kinoshita, K., and Endo, T. (1995). Mitochondrial receptor complex protein. The intermembrane space domain of yeast MAS17 is not essential for its targeting or function. *J Biol Chem* 270:30571–30575.
89. Bolliger, L., Junne, T., Schatz, G., and Lithgow, T. (1995). Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria. *EMBO J* 14:6318–6326.
90. Chacinska, A., Lind, M., Frazier, A.E., Dudek, J., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H.E., Truscott, K.N., Guiard, B., Pfanner, N., and Rehling, P. (2005). Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* 120:817–829.
91. Allen, R., Egan, B., Gabriel, K., Beilharz, T., and Lithgow, T. (2002). A conserved proline residue is present in the transmembrane-spanning domain of Tom7 and other tail-anchored protein subunits of the TOM translocase. *FEBS Lett* 514:347–350.
92. Jansch, L., Kruft, V., Schmitz, U.K., and Braun, H.-P. (1998). Unique composition of the preprotein translocase of the outer mitochondrial membrane from plants. *J Biol Chem* 273:17251–17257.
93. Johnston, A.J., Hoogenraad, J., Dougan, D.A., Truscott, K.N., Yano, M., Mori, M., Hoogenraad, N.J., and Ryan, M.T. (2002). Insertion and assembly of human Tom7 into the preprotein translocase complex of the outer mitochondrial membrane. *J Biol Chem* 277:42197–42204.
94. Honlinger, A., Bomer, U., Alconada, A., Eckerskorn, C., Lottspeich, F., Dietmeier, K., and Pfanner, N. (1996). Tom7 modulates the dynamics of the mitochondrial outer membrane translocase and plays a pathway-related role in protein import. *EMBO J* 15:2125–2137.
95. Sherman, E.L., Go, N.E., and Nargang, F.E. (2005). Functions of the small proteins in the TOM complex of *Neurospora crasssa*. *Mol Biol Cell* 16:4172–4182.
96. Dietmeier, K., Honlinger, A., Bomer, U., Dekker, P.J., Eckerskorn, C., Lottspeich, F., Kubrich, M., and Pfanner, N. (1997). Tom5 functionally links mitochondrial preprotein receptors to the general import pore. *Nature* 388:195–200.

97. Model, K., Meisinger, C., Prinz, T., Wiedemann, N., Truscott, K.N., Pfanner, N., and Ryan, M.T. (2001). Multistep assembly of the protein import channel of the mitochondrial outer membrane. *Nat Struct Biol* 8:361–370.
98. Wiedemann, N., Kozjak, V., Chacinska, A., Schonfisch, B., Rospert, S., Ryan, M.T., Pfanner, N., and Meisinger, C. (2003). Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* 424:565–571.
99. Meisinger, C., Wiedemann, N., Rissler, M., Strub, A., Milenkovic, D., Schonfisch, B., Muller, H., Kozjak, V., and Pfanner, N. (2006). Mitochondrial protein sorting: Differentiation of beta-barrel assembly by Tom7-mediated segregation of Mdm10. *J Biol Chem* 281:22819–22826.
100. Kassenbrock, C.K., Cao, W., and Douglas, M.G. (1993). Genetic and biochemical characterization of ISP6, a small mitochondrial outer membrane protein associated with the protein translocation complex. *EMBO J* 12:3023–3034.
101. Cao, W., and Douglas, M.G. (1995). Biogenesis of ISP6, a small carboxyl-terminal anchored protein of the receptor complex of the mitochondrial outer membrane. *J Biol Chem* 270:5674–5679.
102. Alconada, A., Kubrich, M., Moczko, M., Honlinger, A., and Pfanner, N. (1995). The mitochondrial receptor complex: the small subunit Mom8b/Isp6 supports association of receptors with the general insertion pore and transfer of preproteins. *Mol Cell Biol* 15:6196–6205.
103. Dembowski, M., Kunkele, K.P., Nargang, F.E., Neupert, W., and Rapaport, D. (2001). Assembly of Tom6 and Tom7 into the TOM core complex of *Neurospora crassa*. *J Biol Chem* 276:17679–17685.
104. Stan, T., Ahting, U., Dembowski, M., Kunkele, K.P., Nussberger, S., Neupert, W., and Rapaport, D. (2000). Recognition of preproteins by the isolated TOM complex of mitochondria. *EMBO J* 19:4895–4902.
105. Sollner, T., Griffiths, G., Pfaller, R., Pfanner, N., and Neupert, W. (1989). MOM19, an import receptor for mitochondrial precursor proteins. *Cell* 59:1061–1070.
106. Schneider, H., Sollner, T., Dietmeier, K., Eckerskorn, C., Lottspeich, F., Trulzsch, B., Neupert, W., and Pfanner, N. (1991). Targeting of the master receptor MOM19 to mitochondria. *Science* 254:1659–1662.
107. Ramage, L., Junne, T., Hahne, K., Lithgow, T., and Schatz, G. (1993). Functional cooperation of mitochondrial protein import receptors in yeast. *EMBO J* 12:4115–4123.
108. Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T., and Kohda, D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* 100:551–560.
109. D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem Sci* 28:655–662.
110. Main, E.R., Xiong, Y., Cocco, M.J., D'Andrea, L., and Regan, L. (2003). Design of stable alpha-helical arrays from an idealized TPR motif. *Structure* 11:497–508.
111. Lamb, J.R., Tugendreich, S., and Hieter, P. (1995). Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* 20:257–259.
112. Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21:932–939.
113. Haucke, V., Horst, M., Schatz, G., and Lithgow, T. (1996). The Mas20p and Mas70p subunits of the protein import receptor of yeast mitochondria interact via the tetratricopeptide repeat motif in Mas20p: evidence for a single hetero-oligomeric receptor. *EMBO J* 15:1231–1237.

114. Likic, V.A., Perry, A., Hulett, J., Derby, M., Traven, A., Waller, R.F., Keeling, P.J., Koehler, C.M., Curran, S.P., Gooley, P.R., and Lithgow, T. (2005). Patterns that define the four domains conserved in known and novel isoforms of the protein import receptor Tom20. *J Mol Biol* 347:81–93.
115. Hwa, J.J., Zhu, A.J., Hiller, M.A., Kon, C.Y., Fuller, M.T., and Santel, A. (2004). Germline specific variants of components of the mitochondrial outer membrane import machinery in *Drosophila*. *FEBS Lett* 572:141–146.
116. Mayer, A., Nargang, F.E., Neupert, W., and Lill, R. (1995). MOM22 is a receptor for mitochondrial targeting sequences and cooperates with MOM19. *EMBO J* 14:4204–4211.
117. Haucke, V., Lithgow, T., Rospert, S., Hahne, K., and Schatz, G. (1995). The yeast mitochondrial protein import receptor Mas20p binds precursor proteins through electrostatic interaction with the positively charged presequence. *J Biol Chem* 270:5565–5570.
118. Rapaport, D., Neupert, W., and Lill, R. (1997). Mitochondrial protein import. Tom40 plays a major role in targeting and translocation of preproteins by forming a specific binding site for the presequence. *J Biol Chem* 272:18725–18731.
119. Lithgow, T., and Schatz, G. (1995). Import of the cytochrome oxidase subunit Va precursor into yeast mitochondria is mediated by the outer membrane receptor Mas20p. *J Biol Chem* 270:14267–14269.
120. Lithgow, T., Junne, T., Wachter, C., and Schatz, G. (1994). Yeast mitochondria lacking the two import receptors Mas20p and Mas70p can efficiently and specifically import precursor proteins. *J Biol Chem* 269:15325–15330.
121. Millar, D.G., and Shore, G.C. (1994). Mitochondrial Mas70p signal anchor sequence. Mutations in the transmembrane domain that disrupt dimerization but not targeting or membrane insertion. *J Biol Chem* 269:12229–12232.
122. Suzuki, H., Maeda, M., and Mihara, K. (2002). Characterization of rat TOM70 as a receptor of the preprotein translocase of the mitochondrial outer membrane. *J Cell Sci* 115:1895–1905.
123. Nakai, M., Endo, T., Hase, T., and Matsubara, H. (1993). Intramitochondrial protein sorting. Isolation and characterization of the yeast MSP1 gene which belongs to a novel family of putative ATPases. *J Biol Chem* 268:24262–24269.
124. Wu, Y., and Sha, B. (2006). Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p. *Nat Struct Mol Biol* 13:589–593.
125. Brix, J., Ziegler, G.A., Dietmeier, K., Schneider-Mergener, J., Schulz, G.E., and Pfanner, N. (2000). The mitochondrial import receptor Tom70: identification of a 25 kDa core domain with a specific binding site for preproteins. *J Mol Biol* 303:479–488.
126. Endo, T., and Kohda, D. (2002). Functions of outer membrane receptors in mitochondrial protein import. *Biochim Biophys Acta* 1592:3–14.
127. Sollner, T., Pfaller, R., Griffiths, G., Pfanner, N., and Neupert, W. (1990). A mitochondrial import receptor for the ADP/ATP carrier. *Cell* 62:107–115.
128. Steger, H.F., Sollner, T., Kiebler, M., Dietmeier, K.A., Pfaller, R., Trulzsch, K.S., Tropschug, M., Neupert, W., and Pfanner, N. (1990). Import of ADP/ATP carrier into mitochondria: two receptors act in parallel. *J Cell Biol* 111:2353–2363.
129. Brix, J., Dietmeier, K., and Pfanner, N. (1997). Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *J Biol Chem* 272:20730–20735.

130. Chew, O., Lister, R., Qbadou, S., Heazlewood, J.L., Soll, J., Schleiff, E., Millar, A.H., and Whelan, J. (2004). A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. *FEBS Lett* 557:109–114.
131. Bomer, U., Pfanner, N., and Dietmeier, K. (1996). Identification of a third yeast mitochondrial Tom protein with tetratrico peptide repeats. *FEBS Lett* 382:153–158.
132. Schlossmann, J., Lill, R., Neupert, W., and Court, D.A. (1996). Tom71, a novel homologue of the mitochondrial preprotein receptor Tom70. *J Biol Chem* 271:17890–17895.
133. Horie, C., Suzuki, H., Sakaguchi, M., and Mihara, K. (2003). Targeting and assembly of mitochondrial tail-anchored protein Tom5 to the TOM complex depend on a signal distinct from that of tail-anchored proteins dispersed in the membrane. *J Biol Chem* 278:41462–41471.
134. Hammen, P.K., and Weiner, H. (2000). Structure of the cytosolic domain of TOM5, a mitochondrial import protein. *FEBS Lett* 468:101–104.
135. Schlossmann, J., and Neupert, W. (1995). Assembly of the preprotein receptor MOM72/MAS70 into the protein import complex of the outer membrane of mitochondria. *J Biol Chem* 270:27116–27121.
136. Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J Cell Biol* 164:19–24.
137. Gentle, I.E., Burri, L., and Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. *Mol Microbiol* 58:1216–1225.
138. Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H.E., Guiard, B., Meisinger, C., and Pfanner, N. (2003). An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. *J Biol Chem* 278:48520–48523.
139. Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* 426:862–866.
140. Voulhoux, R., and Tommassen, J. (2004). Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res Microbiol* 155:129–135.
141. Genevrois, S., Steeghs, L., Roholl, P., Letesson, J.J., and van der Ley, P. (2003). The Omp85 protein of Neisseria meningitidis is required for lipid export to the outer membrane. *EMBO J* 22:1780–1789.
142. Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A. (2003). POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem Sci* 28:523–526.
143. Habib, S.J., Waizenegger, T., Lech, M., Neupert, W., and Rapaport, D. (2005). Assembly of the TOB complex of mitochondria. *J Biol Chem* 280:6434–6440.
144. Gratzer, S., Lithgow, T., Bauer, R.E., Lamping, E., Paltauf, F., Kohlwein, S.D., Haucke, V., Junne, T., Schatz, G., and Horst, M. (1995). Mas37p, a novel receptor subunit for protein import into mitochondria. *J Cell Biol* 129:25–34.
145. Ryan, M.T., Muller, H., and Pfanner, N. (1999). Functional staging of ADP/ATP carrier translocation across the outer mitochondrial membrane. *J Biol Chem* 274:20619–20627.
146. Armstrong, L.C., Komiya, T., Bergman, B.E., Mihara, K., and Bornstein, P. (1997). Metaxin is a component of a preprotein import complex in the outer membrane of the mammalian mitochondrion. *J Biol Chem* 272:6510–6518.

147. Milenkovic, D., Kozjak, V., Wiedemann, N., Lohaus, C., Meyer, H.E., Guiard, B., Pfanner, N., and Meisinger, C. (2004). Sam35 of the mitochondrial protein sorting and assembly machinery is a peripheral outer membrane protein essential for cell viability. *J Biol Chem* 279:22781–22785.
148. Ishikawa, D., Yamamoto, H., Tamura, Y., Moritoh, K., and Endo, T. (2004). Two novel proteins in the mitochondrial outer membrane mediate beta-barrel protein assembly. *J Cell Biol* 166:621–627.
149. Armstrong, L.C., Saenz, A.J., and Bornstein, P. (1999). Metaxin 1 interacts with metaxin 2, a novel related protein associated with the mammalian mitochondrial outer membrane. *J Cell Biochem* 74:11–22.
150. Sogo, L.F., and Yaffe, M.P. (1994). Regulation of mitochondrial morphology and inheritance by Mdm10p, a protein of the mitochondrial outer membrane. *J Cell Biol* 126:1361–1373.
151. Boldogh, I.R., Nowakowski, D.W., Yang, H.C., Chung, H., Karmon, S., Royes, P., and Pon, L.A. (2003). A protein complex containing Mdm10p, Mdm12p, and Mmm1p links mitochondrial membranes and DNA to the cytoskeleton-based segregation machinery. *Mol Biol Cell* 14:4618–4627.
152. Meisinger, C., Rissler, M., Chacinska, A., Szklarz, L.K., Milenkovic, D., Kozjak, V., Schonfisch, B., Lohaus, C., Meyer, H.E., Yaffe, M.P., Guiard, B., Wiedemann, N., *et al.* (2004). The mitochondrial morphology protein Mdm10 functions in assembly of the preprotein translocase of the outer membrane. *Dev Cell* 7:61–71.
153. Waizenegger, T., Schmitt, S., Zivkovic, J., Neupert, W., and Rapaport, D. (2005). Mim1, a protein required for the assembly of the TOM complex of mitochondria. *EMBO Rep* 6:57–62.
154. Mnaimneh, S., Davierwala, A.P., Haynes, J., Moffat, J., Peng, W.T., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A., Trochesset, M., Morse, D., *et al.* (2004). Exploration of essential gene functions via titratable promoter alleles. *Cell* 118:31–44.
155. Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M.A., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440:631–636.
156. Boya, P., Roques, B., and Kroemer, G. (2001). New EMBO members' review: viral and bacterial proteins regulating apoptosis at the mitochondrial level. *EMBO J* 20:4325–4331.
157. Nagai, T., Abe, A., and Sasakawa, C. (2005). Targeting of enteropathogenic *Escherichia coli* EspF to host mitochondria is essential for bacterial pathogenesis: critical role of the 16th leucine residue in EspF. *J Biol Chem* 280:2998–3011.
158. Papatheodorou, P., Domanska, G., Oxle, M., Mathieu, J., Selchow, O., Kenny, B., and Rassow, J. (2006). The enteropathogenic *Escherichia coli* (EPEC) Map effector is imported into the mitochondrial matrix by the TOM/Hsp70 system and alters organelle morphology. *Cell Microbiol* 8:677–689.
159. Nougayrede, J.-P., and Donnenberg, M.S. (2004). Enteropathogenic *Escherichia coli* EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cell Microbiol* 6:1097–1111.
160. Muller, A., Rassow, J., Grimm, J., Machuy, N., Meyer, T.F., and Rudel, T. (2002). VDAC and the bacterial porin PorB of Neisseria gonorrhoeae share mitochondrial import pathways. *EMBO J* 21:1916–1929.
161. Massari, P., King, C.A., Ho, A.Y., and Wetzler, L.M. (2003). Neisserial PorB is translocated to the mitochondria of HeLa cells infected with Neisseria meningitidis and protects cells from apoptosis. *Cell Microbiol* 5:99–109.

162. Massari, P., Ram, S., Macleod, H., and Wetzler, L.M. (2003). The role of porins in neisserial pathogenesis and immunity. *Trends Microbiol* 11:87–93.
163. Choi, C.H., Lee, E.Y., Lee, Y.C., Park, T.I., Kim, H.J., Hyun, S.H., Kim, S.A., Lee, S.K., and Lee, J.C. (2005). Outer membrane protein 38 of Acinetobacter baumannii localizes to the mitochondria and induces apoptosis of epithelial cells. *Cell Microbiol* 7:1127–1138.
164. Galmiche, A., Rassow, J., Doye, A., Cagnol, S., Chambard, J.C., Contamin, S., de Thillot, V., Just, I., Ricci, V., Solcia, E., Van Obberghen, E., and Boquet, P. (2000). The N-terminal 34 kDa fragment of *Helicobacter pylori* vacuolating cytotoxin targets mitochondria and induces cytochrome c release. *EMBO J* 19:6361–6370.
165. D'Agostino, D.M., Bernardi, P., Chieco-Bianchi, L., and Ciminale, V. (2005). Mitochondria as functional targets of proteins coded by human tumor viruses. *Adv Cancer Res* 94:87–142.
166. Esaki, M., Shimizu, H., Ono, T., Yamamoto, H., Kanamori, T., Nishikawa, S., and Endo, T. (2004). Mitochondrial protein import. Requirement of presequence elements and TOM components for precursor binding to the TOM complex. *J Biol Chem* 279:45701–45707.

This page intentionally left blank

# 13

# *The Role of the Mia40-Erv1 Disulfide Relay System in Import and Folding of Proteins of the Intermembrane Space of Mitochondria*

JOHANNES M. HERRMANN[a,b] • KARL BIHLMAIER[b] • NIKOLA MESECKE[a]

[a]*Institut für Physiologische Chemie*
*Universität München*
*Butenandtstrasse 5*
*81377 München, Germany*

[b]*Institut für Zellbiologie*
*Universität Kaiserslautern*
*Erwin-Schrödinger-Strasse 13*
*67663 Kaiserslautern, Germany*

## I. Abstract

Many proteins of the mitochondrial intermembrane space (IMS) lack typical mitochondrial targeting signals and are imported on a unique pathway which does neither depend on the hydrolysis of ATP nor on the membrane potential across the inner membrane. These proteins are of small molecular weight and contain characteristic patterns of cysteine residues known as twin $Cx_3C$ or twin $Cx_9C$ motifs. Following their synthesis on cytosolic ribosomes, these proteins traverse the outer membrane through the general translocation pore of the TOM complex. In the IMS,



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25013-9

they are initially bound by disulfide linkages to Mia40, a phylogenetically conserved receptor protein that contains a cysteine-rich domain. By reshuffling of the disulfide bonds, the imported proteins are released from the Mia40 receptors in a folded conformation. According to the folding trap hypothesis, the stably folded structure thereby prevents their back-translocation into the cytosol and maintains the proteins permanently in the IMS. For the formation of the disulfide bridges, the IMS contains a conserved sulfhydryl oxidase named Erv1. This flavin adenine dinucleotide (FAD)-binding enzyme interacts with Mia40 and converts it into its oxidized, functionally active state. The Erv1 sulfhydryl oxidase together with the Mia40 receptor thereby functions as a disulfide relay system in mitochondria, which presumably provides the net energy for the translocation of proteins into the IMS. This chapter summarizes our knowledge on Mia40 and Erv1 and discusses a hypothetical model on the molecular mechanism by which these proteins facilitate the import of proteins into the IMS of mitochondria.

## II. Introduction

The intermembrane space (IMS) of mitochondria is enclosed by the outer and the inner membrane of the organelle. Since both mitochondrial membranes are only a few nanometers apart, the IMS is a rather tiny compartment of the cell. Nevertheless, components of the IMS fulfill a variety of crucial functions in different processes, like in the transport of proteins, electrons, or metal ions, in the assembly of inner membrane proteins, in cellular respiration, and other metabolic processes. In addition, several apoptotic components are sequestered in the IMS until their release triggers the programmed cell death.

The IMS can be subdivided into two distinct subcompartments, the intracristae space and the lumen between the outer and the inner membrane, which is also called the ''external IMS''. Both subcompartments are separated by cristae junctions, which presumably form rather tight openings at the necks of the cristae [1, 2]. We know very little about the physicochemical properties of both subcompartments. Due to the presence of porins in the outer membrane, molecules up to a molecular mass of about 2–6 kDa can freely diffuse into the external IMS [3]. The concentration of glutathione in the external IMS is therefore expected to be similar to that of the cytosol, suggesting that the IMS is a strongly reducing environment. However, a number of disulfide bridges have been reported in proteins of the IMS suggesting that, despite its high concentration of reduced glutathione, cysteine residues can be actively oxidized (see Table 13.1). Whether

TABLE 13.1

PROTEINS FOR WHICH DISULFIDE BONDS IN THE IMS HAVE BEEN REPORTED

| Protein | References | Function |
|---|---|---|
| CCS | [16, 74] | Copper chaperone of Sod1 |
| Cox11 | [75] | Biogenesis of cytochrome oxidase |
| Cox12 | [69, 76] | Biogenesis of cytochrome oxidase |
| Cox17 | [69, 72] | Biogenesis of cytochrome oxidase |
| Cox19 | [69, 77] | Biogenesis of cytochrome oxidase |
| Cox23 | [69, 78] | Biogenesis of cytochrome oxidase |
| Erv1 | [79] | Sulfhydryl oxidase |
| Mia40 | [24–25, 49] | Import receptor |
| Rieske protein | [80] | Subunit of cytochrome reductase |
| Qcr6 | [80] | Subunit of cytochrome reductase |
| small Tim proteins | [61–63] | Components of import machinery |
| Sco1 | [81] | Biogenesis of cytochrome oxidase |
| Sod1 | [16, 82] | Superoxide dismutase |

this oxidation occurs both in the external IMS and the intracristae space, or mainly in the latter, remains unclear.

## III. Protein Import Routes into the IMS

All proteins of the IMS are encoded by nuclear genes and synthesized in the cytosol. However, IMS proteins often lack classical mitochondrial presequences, and their import into the organelle is mediated by pathways distinct from the import route that directs proteins into the matrix of mitochondria. While matrix proteins appear to be imported consistently on one general route, proteins of the IMS use different import mechanisms (for overview see [4–6]). On the basis of their targeting signals and their energetic requirements, three classes of IMS proteins can be distinguished (Figure 13.1):

- Class I proteins contain so-called bipartite presequences at their very N-termini consisting of a mitochondrial targeting sequence followed by a hydrophobic sorting domain. Protein translocases in both mitochondrial membranes recognize the mitochondrial targeting signal and transport it into the matrix in an ATP- and membrane potential-dependent reaction. The hydrophobic sorting domain then leads to a translocation arrest at the level of the inner membrane and mediates the lateral insertion of the preprotein. Upon proteolytic cleavage, the matured proteins are released into the IMS. A number of larger proteins fall into this class

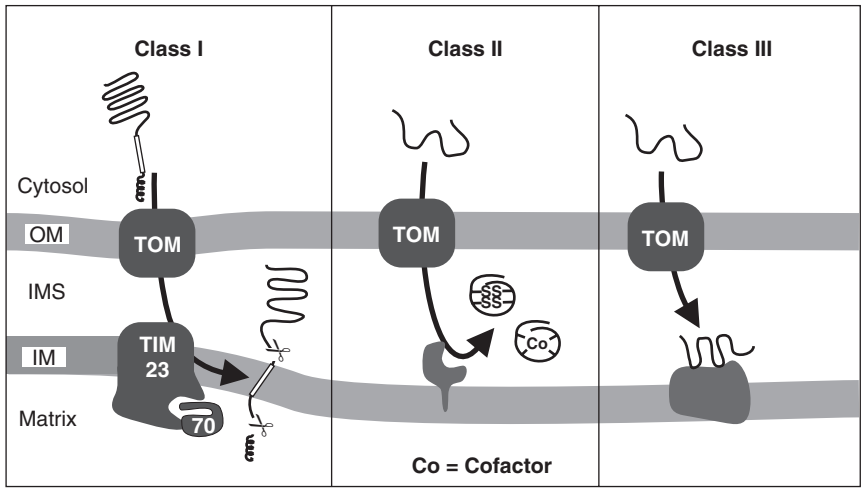



FIG. 13.1. Proteins of the IMS can be subdivided into three classes on the basis of their import mechanisms. Class I proteins contain bipartite presequences consisting of a matrix-targeting signals (shown as a helix) followed by hydrophobic sorting domains (shown as white box). Following import, the hydrophobic domains integrate into the inner membrane. Cleavage by processing peptidases finally releases the proteins into the IMS. Class II proteins traverse the TOM complex and transiently interact with receptors in the IMS. They are then released from the receptors in a folded conformation. Thereby folding can be stabilized by the binding of cofactors (Co) or by the oxidation of cysteine residues. Class III proteins are translocated through the general import pore in the TOM complex into the IMS where they permanently associate with binding sites on the inner or the outer membrane.

like apoptosis-inducing factor, cytochrome *c* peroxidase, cytochrome $b_2$, endonuclease G, or Smac/Diablo [4, 7–9].

- Class II proteins are usually of small molecular weight and comprise only one folding domain. The folded conformation is typically stabilized either by binding of cofactors like heme or metal ions or by the formation of intramolecular disulfide bridges in these proteins. Cytochrome *c* was the first representative of this group for which the import process was studied in detail [10–15]. Apocytochrome *c* can cross the outer membrane through the TOM complex in both directions. In the IMS, apocytochrome *c* binds to cytochrome *c* heme lyase. This enzyme incorporates a heme group into apocytochrome *c* before it releases the holocytochrome *c* as soluble stably folded protein into the IMS that is unable to cross the TOM complex. Cytochrome *c* heme lyase functions both as a receptor and as a converting enzyme that catalyzes the stable folding of cytochrome *c*. A similar mechanism was found also for other IMS proteins like for Cu/Zn-superoxide dismutase which transiently interacts with the copper chaperone Ccs1 [16]

or for a variety of small cysteine-containing proteins which transiently interact with the Mia40 receptor protein before they are released in a stably folded conformation into the IMS. The import of superoxide dismutase and the cysteine-containing proteins depends on the formation of intramolecular disulfide bridges and relies on the presence of a conserved sulfhydryl oxidase in the IMS, named Erv1. The Mia40/Erv1-dependent import process will be described in detail in this chapter.

- Class III proteins represent not soluble factors of the IMS but rather are permanently associated with binding sites at the outer or inner membrane. The affinity to these binding sites presumably provides the energy that drives the import reaction. Representatives of this class are cytochrome *c* heme lyases [17] or creatine kinase [18].

## IV. Mia40, an Import Receptor in the IMS

### A. Structural Organization of Mia40 Proteins

Mia40 is an essential mitochondrial protein that is conserved between fungi, plants, and animals including humans. Mia40 was initially identified in mitochondria of baker's yeast [19] and what we know about this protein is almost entirely based on studies of the yeast homologue [20–22]. All Mia40 homologues share a highly conserved C-terminal domain containing six invariant cysteine residues (Figure 13.2A). Fungal members of the

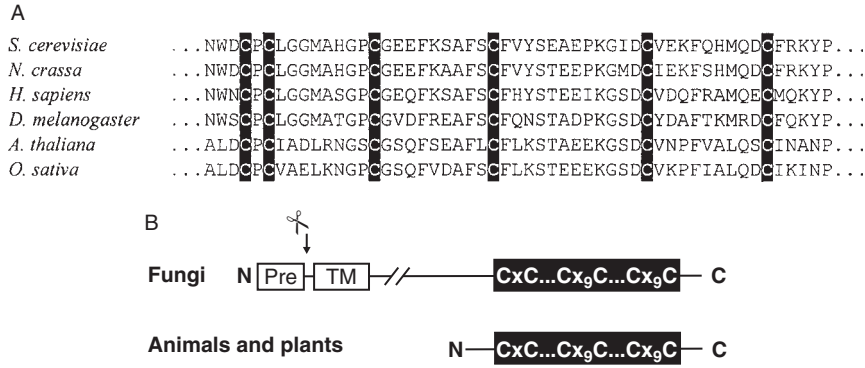



FIG. 13.2. Structure of Mia40. (A) Sequence alignment of the cysteine-rich domain of Mia40 proteins of the species indicated. Invariant cysteine residues are indicated. (B) Schematic representation of the organization of Mia40 proteins. The black box depicts the conserved cysteine-rich domain of Mia40. For the fungal protein, the processing site for the mitochondrial presequence peptidase (MPP) is indicated. Pre, mitochondrial presequence and TM, transmembrane domain.

Mia40 family, but not those of animals or plants, are synthesized with N-terminal mitochondrial targeting signals followed by hydrophobic membrane anchor domains (Figure 13.2B). While it was initially proposed that this region represents a bipartite presequence that targets the protein into the IMS [20], more recent studies suggested that the hydrophobic region remains present on the mature protein and anchors it permanently to the inner membrane [21, 22]. The stable association of Mia40 with the inner membrane is, however, not essential for its function; Mia40 mutants in which the conserved C-terminal domain was fused to the bipartite presequence of cytochrome $b_2$ were still viable [22].

In *Saccharomyces cerevisiae*, but not in other fungi, the hydrophobic membrane anchor and the conserved domain are separated by a highly acidic region of unknown function. Interestingly, alignments of the DNA sequence of this region with that of the *S. cerevisiae* genome revealed a region on chromosome XII (nucleotides 351207–351591) of 58% identity. Hence, a duplication of this DNA segment during the evolution of baker's yeast may have resulted in an insertion within the *MIA40* gene and thereby created this unique acidic domain.

The conserved domain of Mia40 comprises roughly 60 amino acid residues. It contains six cysteine residues present in the pattern CPC-$x_8$-C$x_9$C-$x_{12}$-C$x_9$C (Figure 13.2). The function of the individual cysteine residues is not known. Exchange of the first, middle, or last pair of cysteine residues by serine residues is lethal [21, 22]. The pattern of the latter two pairs of cysteine residues resembles the twin C$x_9$C motif found in many Mia40 substrate proteins (see below). Why this signature is present in both Mia40 and its substrate proteins is not clear.

Recombinant Mia40 was shown to be able to complex metal ions like copper and zinc [21]. Moreover, the removal of metal ions with chelating reagents led to a proteolytic instability of Mia40. The cysteine residues were crucial for the binding of metal ions suggesting that they are coordinated by the cysteine motif in Mia40 [21]. However, *in vivo*, some of the cysteine residues in Mia40 appear to be predominantly present in an oxidized state forming intramolecular disulfide bonds [23–25]. Which of the cysteine residues take part in the formation of the disulfide bridges is not known. *In vivo*, Mia40 might cycle between a reduced, metal-bound, and an oxidized disulfide-bridged conformation and, hence, both conformations might represent physiologically relevant states (see below).

## B. The Function of Mia40 in the IMS

Mia40 is an essential protein and its deletion is lethal on both fermentable and nonfermentable carbon sources [20–22]. The depletion of endogenous Mia40 *in vivo* leads to a concomitant loss of proteins with twin C$x_3$C

and twin $Cx_9C$ motifs from the IMS. Import experiments into isolated mitochondria indicate that Mia40 directly binds to these proteins in a transient manner after their translocation across the outer membrane. In the absence of Mia40, proteins with twin $Cx_3C$ and twin $Cx_9C$ motifs fail to be efficiently imported into mitochondria, whereas other IMS proteins like cytochrome *c* or cytochrome *c* heme lyase are not affected. This suggests that Mia40 has a receptor-like function for the import of IMS proteins that contain conserved patterns of cysteine residues. It is presently not clear whether, in addition to its role as a receptor, Mia40 is also more actively involved in the folding or assembly of IMS proteins.

Human Mia40 is significantly smaller than the yeast protein (17 versus 40 kDa) and consists almost exclusively of the conserved domain that shows 75% sequence similarity to the yeast homologue [23]. Due to the lack of the membrane-anchoring region, the human Mia40 forms a soluble factor in the IMS that, similar to its fungal homologue, forms reduced and oxidized conformers. Like in yeast, depletion of human Mia40 leads to reduced levels of IMS proteins with twin $Cx_3C$ and twin $Cx_9C$ signatures and finally causes cell death [23]. Thus, despite the differences in their structural organization, the Mia40 proteins of yeast and human appear to exhibit comparable functions.

## V. Erv1, a Disulfide Oxidase in the IMS

The Erv1 protein of *S. cerevisiae* was initially identified as a protein *e*ssential for *r*espiration and *v*egetative growth [26]. Yeast Erv1 was the first member to be discovered of a large family of proteins which are widely distributed in eukaryotes [27–29]. These proteins share a conserved flavin adenine dinucleotide (FAD)-binding domain that contains a sequence of two cysteine residues spaced by two amino acid residues, the CxxC motif (see below). It is characteristic for redox active proteins and a sulfhydryl oxidase activity could be assigned to members of this family [30–33]. The best-characterized member of the family is Erv2, a sulfhydryl oxidase of the endoplasmic reticulum (ER) of fungi [31, 34], and most of what we know about the molecular activity of this group of sulfhydryl oxidases is derived from studies on Erv2.

### A. Structural Organization of Erv1 Proteins

Erv1 proteins of fungi and animals show a consistent architecture comprising two structural segments (Figure 13.3). The N-terminal segment that in *S. cerevisiae* consists of 72 amino acid residues (8.8 kDa) is rich in glycine and proline residues and presumably represents a poorly structured flexible region. This segment of Erv1 is characterized by the presence of a CxxC

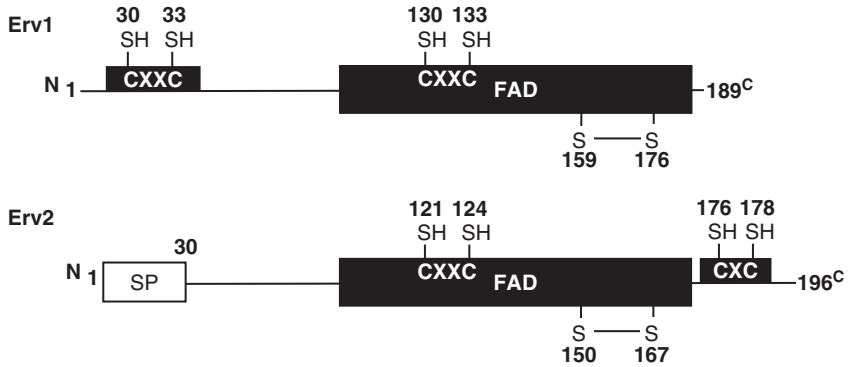



FIG. 13.3. Structural organization of Erv1 and Erv2. The schematic architecture of Erv1 and Erv2 of *S. cerevisiae* is sketched. The numbers indicate the positions of specific amino acid residues in the proteins. The conserved FAD-binding domains are shown as black boxes. The active CxxC motifs and the structural disulfide bridges in these domains are indicated. SP, signal peptide for ER targeting of Erv2. See text for details.

motif that is essential for Erv1 function [35]. The C-terminal segment forms an FAD-binding domain [36, 37] that in *S. cerevisiae* consists of 117 amino acid residues (12.8 kDa). This domain is well conserved throughout the Erv1/Erv2 protein family; the sequence identity between Erv1 and Erv2 proteins of this region is in the range of about 30% and functionally relevant residues are largely identical between both proteins. Erv1 and Erv2 form homodimers (Figure 13.4A) which, depending on the redox state, can be transiently linked by disulfide bonds [30, 36, 38]. Achievements in the crystallization of the C-terminal domain of Erv2 provided a detailed structure of the FAD-binding domain with 1.5-Å resolution which allowed significant insights into the molecular mechanism by which this group of redox enzymes can introduce disulfide bonds into substrate proteins [36].

Purification of recombinant Erv1 with subsequent absorbance spectroscopy revealed an FAD cofactor noncovalently bound to the C-terminal fragment [30]. Its presence in Erv2 was demonstrated spectroscopically [31] as well as by crystal structure determination [36] in an unusual horseshoe-like conformation with the flavin and isoalloxazine parts buried in the interior of the domain and the polar ribose and phosphate groups exposed to the solvent [36, 39]. The FAD group and aromatic side chains form unusual planar-stacking interactions (Figure 13.4A). The amino acid residues that form the FAD-binding site are highly conserved between Erv1 and Erv2 (Figure 13.4B). In both proteins, the FAD-binding fold is stabilized by a conserved permanent disulfide bond (Figure 13.3B, indicated with s).

The essential CxxC motif of the FAD-binding domain is in proximity to the isoalloxazine ring, facilitating an efficient electron transfer to FAD. In general, CxxC motifs are a hallmark for redox active proteins and usually present in a structure called the thioredoxin (TRX) fold. Although Erv1 and Erv2 do not share sequence homology with Ero1, the tertiary structure of the FAD-binding domain of these sulfhydryl oxidases is surprisingly similar [40].

While the FAD-binding domains of Erv1 and Erv2 proteins are very similar, both proteins differ in the sequences that flank this region. Erv2 proteins contain a characteristic cysteine-glycine-cysteine motif C-terminally to the FAD-binding domain that is absent in Erv1 proteins. This motif is part of a mobil lever arm that interacts with the CxxC motif in the FAD-binding domain of the opposing Erv2 subunit in the dimeric Erv2 complex via a transient disulfide bridge and is thereby oxidized and activated. Erv2 can then interact with substrate proteins to introduce disulfide bridges [36, 38]. In Erv1 proteins, the N-terminal CxxC motif might play an analogous function in the communication between the FAD-binding domain and substrate proteins. This CxxC motif is connected to the FAD-binding domain via a stretch of variable length that contains a high number of helix-breaking residues; for example, in *S. cerevisiae*, 19.4% of the residues in this stretch are proline and glycine residues. While such a lever armlike function of the N-terminal part of Erv1 is intriguing, experimental evidence for such a role of this region is still missing.

## B. OTHER PROTEINS WITH ERV1-LIKE DOMAINS

Little is known on the molecular function of the human Erv1 homologue, which was also named *a*ugmenter of *l*iver *r*egeneration (ALR). ALR was initially identified in a screen for hepatotrophic growth factors [27]: to select growth factors for hepatocytes, extracts from hepatoectomized rat livers were injected into rats in which a part of the liver had been removed. Thereby human Erv1 was isolated as a hepatic stimulator substance, HSS, and termed hepatopoietin, HPO (for review see [29, 41]). However, the mechanism by which human Erv1 stimulates the growth of hepatocytes is still elusive and it is even not clear whether, *in vivo*, human Erv1 plays a physiologically relevant role in liver development.

Apart from Erv1/ALR and Erv2, two additional groups of sulfhydryl oxidases exist that contain Erv1-like FAD-binding domains, namely members of the Quiescin/sulfhydryl oxidase (QSOX) family and viral Erv1-like proteins (for review see [28]).

The QSOX proteins presumably arose from an ancient fusion event that combined a TRX domain including a classical CxxC motif to an Erv1-like

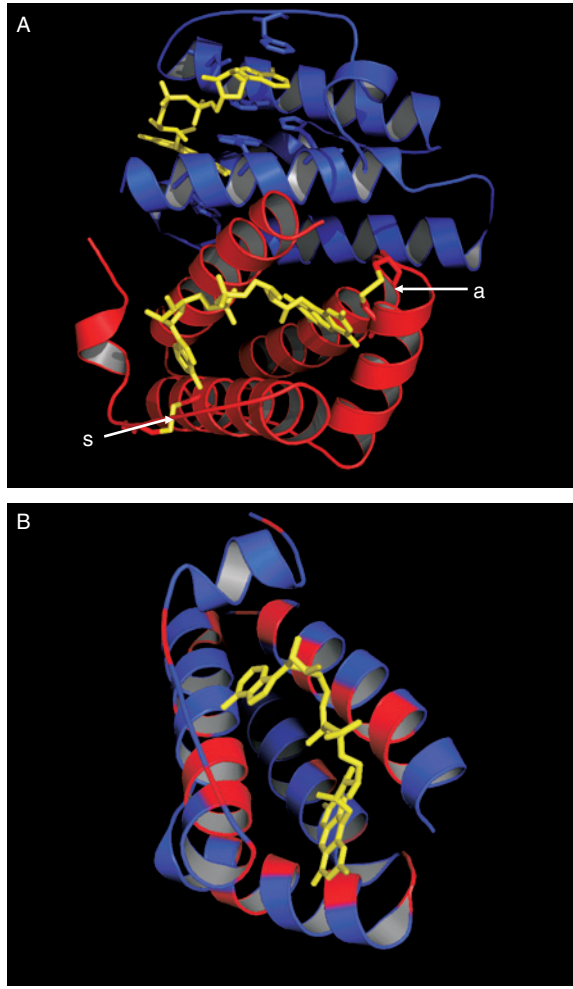



FIG. 13.4. Comparison of the molecular structures of Erv1 and Erv2. A colored version of this figure is displayed at the end of this volume. (A) Shown is a model of the molecular structure of the FAD-binding domains of an Erv1–Erv1 homodimer generated using the PyMOL software by molecular replacement on the basis of the PDB file 1JRA of Erv2 [36]. The FAD cofactors are shown in light gray (yellow in the colored version). Note that different structural details are indicated in the upper and lower subunit, which *in vivo* are identical. In the upper subunit (shown in blue in the colored version), the aromatic side chains that fix the FAD group via hydrophobic stacking interactions are illustrated. In the lower subunit (shown in red in the colored version), the two disulfide bridges in the FAD-binding domain are depicted in light gray (or yellow, respectively). The redox-active disulfide bridge (indicated by a) is in direct proximity to the isoalloxazine group of FAD and exposed to the surface of Erv1. The structural disulfide bridge (indicated by s) fixes a flexible unstructured C-terminus of

sulfhydryl oxidase domain [42]. QSOX proteins are present in animals and mainly distributed in compartments of the secretory pathway. These proteins function as sulfhydryl oxidases in which the N-terminal TRX domain might play a role in the transfer of the electrons between the FAD-binding domain and substrate proteins. Thus, the mode of function is reminiscent to that of Erv1 and Erv2, where also two functional pairs of cysteine residues are believed to pass on electrons in a sequential transfer reaction.

Proteins of QSOX-like organization are present in plants (Figure 13.5). Moreover, green algae like *Chlamydomonas* and cyanobacteria also contain proteins containing a domain of significant similarity to the FAD-binding domain of Erv1. These proteins, however, lack the TRX-like domain.

Interestingly, even some viral genomes encode Erv1 homologues. A number of DNA viruses produce stably folded capsid proteins in the cytosol by introduction of disulfide bonds. In order to oxidize cytosolic cysteine residues, these viruses encode Erv1-like sulfhydryl oxidases which are expressed in the cytosol of the host cell [43–45]. Well-studied examples for such viruses are the vaccinia virus and the *African swine fever virus*.

## C. FUNCTIONS OF ERV1 IN THE IMS

Erv1 is located in the IMS of mitochondria [46]. Although the protein was originally reported to be present also in the cytosol [35], convincing evidence for a location of Erv1 outside of mitochondria is still missing.

A yeast strain expressing a temperature-sensitive allele of *ERV1* was initially identified that, on shift to restrictive growth conditions, showed a variety of pleiotropic defects. Based on these observations, it was suggested that Erv1 plays an essential role in various cellular processes ranging from the generation of a functional respiratory chain to a role in the distribution of mitochondria within the cell [26, 47, 48].

So far, only one direct substrate protein of Erv1 was identified which is the Mia40 import receptor in the IMS (see below). Whether the variety of defects observed in *erv1* mutants is indirectly caused by defects in the import process of different Mia40-dependent IMS proteins, or whether Erv1 introduces disulfide bridges in several different substrate proteins, is unclear.

Erv1 to the α-helix that is close to the adenine group of the FAD molecule. (B) Enlarged view of the FAD-binding pocket of Erv2 [36]. Residues shown in light gray (or red in the colored version) are identical between Erv1 and Erv2, whereas residues in dark gray (or blue, respectively) are different. This illustrates that the side chains that contribute to the binding of FAD are practically identical between Erv1 and Erv2.

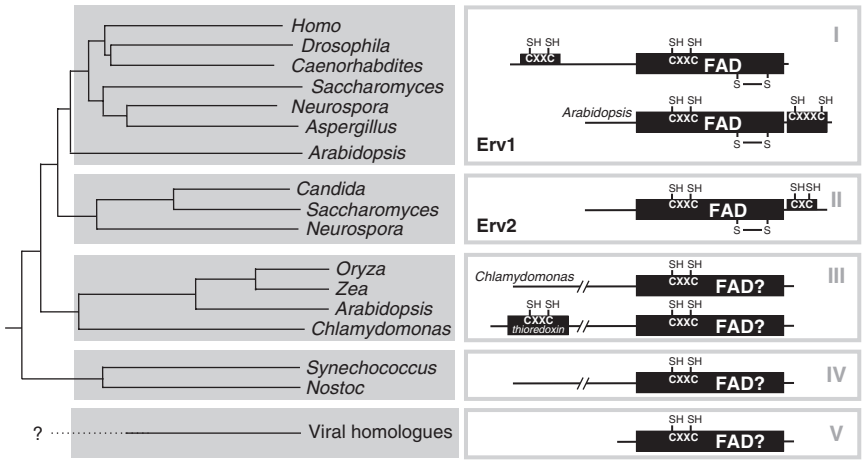



FIG. 13.5. Phylogenetic analysis of members of the Erv1 family. Proteins with homology to Erv1 were aligned and a phylogenetic tree was calculated using the DNAman software package (Lynnon Corporation, Quebec, Canada). The structural organization of members of the different branches is sketched as in Figure 13.2. Members of the QSOX family were not included in the alignment, but their structural organization is similar to that of branch III. The viral Erv1 proteins are highly variable and their position in the tree can therefore not be reliably traced. The sequences used for the analysis are for group I *Homo sapiens* (NP_005253), *Drosophila melanogaster* (AY094854.1), *Caenorhabditis elegans* (AAB97554), *S. cerevisiae* (NP_075527), *Neurospora crassa* (XP_959716), *Aspergillus nidulans* (XP_660631), *Arabidopsis thaliana* (AAM63908), for group II *Candida albicans* (XP_720875), *S. cerevisiae* (NP_015362), *N. crassa* (NP_015362), for group III *Oryza sativa* (AAT85195), *Zea mays* (AAW66880), *A. thaliana* (AAF31025), *Chlamydomonas reinhardtii* (AAV32452), for group IV *Synechococcus* sp. *CC9605* (ABB34539), *Nostoc punctiforme* (ZP_00107616), and for group V *African swine fever virus* (NP_042767).

## *1. Erv1 Is Critical for the Import of Certain Proteins into the IMS*

It was shown that Erv1 is directly involved in the translocation of proteins that contain characteristic twin $Cx_3C$ and $Cx_9C$ motifs from the cytosol to the IMS. In this process, Erv1 is essential for the formation of disulfide bonds in the Mia40 receptor in the IMS converting or maintaining this receptor protein into its active form [24, 25, 49]. In addition, depletion of Erv1 leads to the loss of Cu/Zn-superoxide dismutase from the IMS, suggesting that the import of this protein also depends on Erv1 function. The role of Erv1 in the biogenesis of twin $Cx_3C$ and twin $Cx_9C$ proteins will be described in detail in the following sections.

### 2. *Erv1 Is Critical for the Assembly of Iron–Sulfur Clusters into Cytosolic Proteins*

Yeast Erv1 was identified to be necessary for the maturation of cytosolic but not of mitochondrial iron–sulfur proteins [46]. The generation of iron–sulfur clusters for both mitochondrial and cytosolic proteins depends on a series of biosynthetic reactions in the mitochondrial matrix [50]. A so far unidentified precursor form of these clusters appears to be synthesized in the matrix before it is transported via an ABC transporter into the IMS and further into the cytosol where it is utilized for the biosynthesis of iron–sulfur proteins [50]. Erv1 was shown to be critical specifically for the formation of cytosolic but not of mitochondrial iron–sulfur proteins, suggesting that Erv1 plays a critical role in the export of the precursor form of iron–sulfur clusters. Defects in the transport of iron sulfur clusters into the cytosol appear shortly after loss of Erv1 function [46], leading to the hypothesis that Erv1 is directly involved in this process.

### 3. *Erv1 Is Critical for Normal Mitochondrial Morphology and Distribution*

Mutants in Erv1 were reported to be unable to distribute mitochondria properly within the cell. Instead of a normal mitochondrial network, *erv1* mutants show enlarged mitochondrial structures that lack cristae membranes and that are preferentially localized at two poles of the cell [47]. The molecular basis of these morphology defects is elusive but it appears likely that they are indirectly caused by the compromised import of morphogenic components into the IMS. Indeed, one of the putative Erv1 substrates, Mdm35, was initially identified as a component critical for mitochondrial distribution and morphology [51].

## VI. A Model of Mia40-Erv1-Mediated Import

Mainly on the basis of *in vitro* import experiments, a model was proposed of how Erv1 and Mia40 mediate the import of proteins into the IMS [25, 52]. It should be stressed that this model is still rather hypothetical, but explains most of the observed results (Figure 13.6).

Proteins of the Mia40 pathway use the general import pore of the TOM complex to cross the mitochondrial outer membrane. However, in contrast to proteins with typical mitochondrial targeting signals, they apparently do not detectably interact with surface receptors on the mitochondria [53]. It is unclear why these proteins appear to avoid these high-affinity interactions,

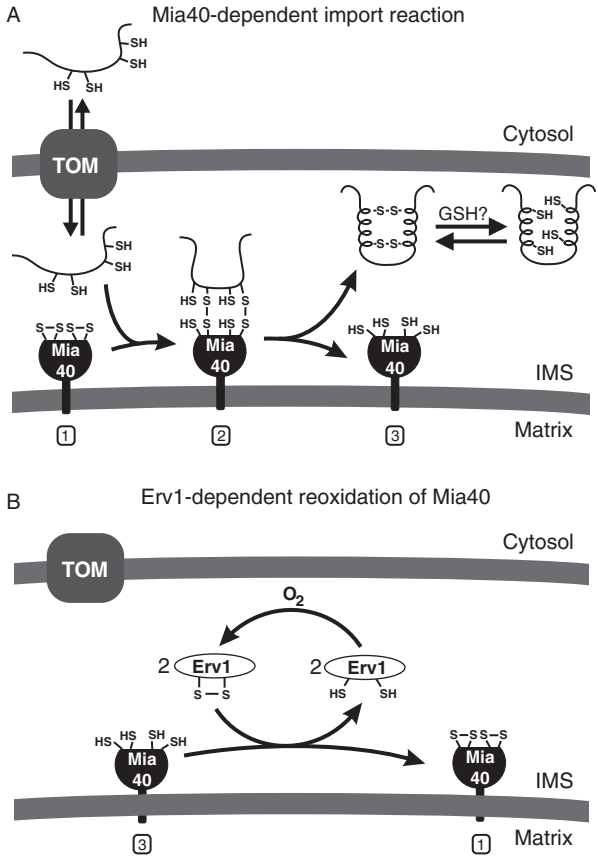



FIG. 13.6. Model of the Mia40-Erv1-dependent protein import. (A) Stage 1: Mia40 is present in mitochondria in its oxidized, active form. This species is able to interact with a newly imported substrate protein that enters the IMS in a reduced conformation via the pore of the TOM complex. Stage 2: Mia40 forms intermolecular disulfide bonds with its substrate protein leading to an arrest of the substrate in the IMS. Stage 3: the substrate is released from Mia40 in the oxidized and folded conformation. Since, in its folded state, these proteins are unable to traverse the TOM complex, this leads to their stable trapping in the IMS. After the release, Mia40 is present in the reduced, inactive state. (B) Reduced Mia40 is reoxidized by the sulfhydryl oxidase Erv1 and therefore regenerated for the next round of import (see text for further details).

but it seems conceivable that the Mia40-dependent import process does not provide the energy required to release the proteins from these binding sites since this pathway functions independent of potent sources of external energy like the hydrolysis of ATP or the membrane potential. Thus, instead of being directed by a series of interactions of increasing affinity, these

proteins may initially reach the IMS by a diffusion-like process and are then trapped at the *cis* side of the TOM complex by binding to Mia40. In order to bind these proteins, Mia40 needs to be present in its active, that is (at least partially), oxidized form (Figure 13.6, stage 1). It was proposed that the binding of the substrates occurs by reorganization of the initially intramolecular disulfide bonds in Mia40 to intermolecular bridges between Mia40 and its substrates (Figure 13.6, stage 2).

The substrate proteins are subsequently released from Mia40, presumably in an oxidized state. The imported proteins are locked in a stably folded conformation and thereby irreversibly trapped in the IMS as the folded proteins cannot pass the protein-conducting channel of the TOM complex. According to this folding trap model, the conversion of the initially unfolded protein into a stably folded structure determines the directivity of the translocation process [53, 54]. The release of the substrate from Mia40 was proposed to leave Mia40 in a reduced, nonfunctional form (Figure 13.6, stage 3). However, experimental data on the redox states of Mia40 and the released substrate proteins are still lacking. But since the sulfhydryl oxidase activity of Erv1 is required to regenerate and reactivate Mia40 (Figure 13.6B) [25], it appears likely that electrons are passed from the substrate proteins via Mia40 to Erv1. Erv1 then presumably transfers the electrons to molecular oxygen, explaining the observed import defects of IMS proteins under low oxygen conditions [25]. In general, FAD-dependent sulfhydryl oxidases can directly use molecular oxygen as final electron acceptor producing $H_2O_2$. It was proposed that *in vivo*, Erv1 can use cytochrome *c* as electron acceptor, which then transfers the electrons via cytochrome oxidase to oxygen [24]. Experiments *in vitro* indeed revealed a direct interaction of human Erv1 and cytochrome *c* where cytochrome *c* was a 100-fold better electron acceptor for Erv1 than molecular oxygen [55]. In contrast to Erv1 and Mia40, cytochrome *c* is, however, not essential, indicating that cytochrome *c* cannot be the exclusive electron acceptor of Erv1.

In addition to its role in protein translocation, Erv1 might also be involved in the assembly of small Tim proteins [24, 49]. Experiments with *erv1* mutants revealed specific defects in the formation of TIM10 complexes from Tim9 and Tim10 subunits [49]. In addition, some temperature-sensitive *mia40* mutants still allow the binding of small Tim proteins to Mia40 but block their subsequent assembly [20]. These data indicate a role, either direct or indirect, of Erv1 in the assembly of Tim complexes that cannot be explained by the Erv1-dependent reoxidation of reduced Mia40 suggesting that the function of Erv1 in the biogenesis of IMS proteins is not restricted to protein import.

## VII. Substrate Proteins of the Mia40-Erv1 Pathway

### A. Proteins of the Twin $Cx_3C$ Family

The best-studied substrates of the Mia40 pathway are members of the small Tim proteins. These proteins are characterized by the presence of an invariant pattern of four cysteine residues forming two $Cx_3C$ motifs, and therefore often are called twin $Cx_3C$ proteins. Mitochondria of fungi, animals, and plants contain typically five members of this family which in yeast were named Tim8, Tim9, Tim10, Tim12, and Tim13 (for review see [56, 57]). They form hexameric complexes in the IMS that play an essential role in the transfer of hydrophobic carrier proteins from the TOM complex to their insertion sites in the inner membrane.

The presence of each of the four cysteine residues in this twin $Cx_3C$ motif is critical for the efficient import of small Tim proteins into the IMS as well as for their stable folding and assembly [53, 58, 59]. The crystal structure of the TIM9–TIM10 complex was solved, suggesting that the twin $Cx_3C$ motif forms a hairpin-like structure stabilized by intramolecular disulfide bridges, one between the two central and one between the two distal cysteine residues [60]. This central core in the small Tim proteins is flanked by less-structured regions, which might participate in the binding of hydrophobic carrier proteins. An oxidized state of the four cysteine residues is supported by several studies both *in vitro* and *in vivo* [54, 61–63]. On the other hand, experiments on purified small Tim proteins showed that the twin $Cx_3C$ motif has the ability to bind zinc ions [58, 64, 65] and the presence of reduced Tim13 was shown *in vivo* [53]. It was proposed that this seeming contradiction between the redox states of endogenous small Tim proteins might be due to an alteration between reduced and oxidized states, potentially in dependence of the redox conditions in the cell or the specific functional state in which the proteins were analyzed [4, 56, 66, 67].

It was recently suggested that the binding of zinc to reduced small Tim proteins occurs already in the cytosol right after their synthesis. According to this hypothesis, the zinc ions stabilize small Tim proteins in a reduced, import-competent conformation. Once translocated to the IMS, the zinc ions are released from the small Tim proteins, presumably during the Mia40-dependent oxidation of the proteins [68].

### B. Proteins of the Twin $Cx_9C$ Family

The second group of Mia40-dependent substrate proteins is characterized by the presence of two pairs of cysteine residues, each spaced by short helices formed by nine variable residues. These helices interact in an

antiparallel orientation and are stabilized by intramolecular disulfide bridges between the flanking cysteine residues [69]. Examples of proteins that contain twin $Cx_9C$ motifs are: Cox17, Cox19, Cox23, all of which are factors required for the assembly of cytochrome oxidase; Som1, a subunit of the Imp1 protease; the FeS protein 5 of complex I and Mdm35, a protein involved in mitochondrial morphogenesis. The best-characterized member of this group is the Cox17 protein, a soluble copper chaperone of the IMS [70, 71]. In *S. cerevisiae*, Cox17 contains 69 amino acid residues, 6 of these representing conserved cysteine residues. The analyses on the Cox17 structure [69, 72, 73] suggested that this protein can be present in up to three conformers which differ in their redox states and in the number of copper-binding sites: (1) the species in which the cysteine residues of the twin $Cx_9C$ motif form two disulfide bridges (residues C26–C57 and C36–C47) does not bind copper. (2) On isomerization of the C26–C57 bond to a C24–C57 bond, Cox17 forms a species that coordinates one copper ion. (3) The completely reduced form of Cox17 is able to bind up to four copper ions per Cox17 monomer. Whether all these conformers represent physiological states of the protein and whether changes in the redox states of Cox17 are relevant for its function in the binding and delivery of copper to other proteins are unclear. However, the different redox states found in Cox17 are reminiscent of those reported for small Tim proteins and it appears conceivable that both groups of IMS proteins shuttle between oxidized and reduced, metal-bound states *in vivo*.

## VIII. Perspectives

The recent discovery of a disulfide relay system of Mia40 and Erv1 in the IMS of mitochondria revealed insights into a completely novel and exciting intracellular protein translocation pathway. The oxidation of cysteine residues obviously shows parallels to processes in the bacterial periplasm and in the ER. However, some features are strikingly different between these systems. For example, unlike in the periplasm and in the ER, mitochondrial proteins typically show four conserved cysteine residues forming characteristic twin $Cx_3C$ and twin $Cx_9C$ motifs. These patterns allow, on binding of zinc or copper ions, the stabilization of the structural fold even when the cysteine residues are reduced [68]. It appears possible that this specific property is used to encounter the specific necessity in the IMS to retain proteins folded in a reducing environment. Future studies will have to address whether endogenous IMS proteins indeed shuttle between oxidized and reduced, metal-bound conformers and, if so, which physiological relevance such conformational switches have. Since the oxidation by Erv1 depends on

oxygen, it may be speculated that this system is used to adapt mitochondrial functions to the respective oxygen concentrations in the mitochondria.

## Acknowledgments

We thank Michael Groll for help with the modeling of the Erv1 structure, and Walter Neupert, Kai Hell, Soledad Funes, and Melanie Bien for comments on the chapter. This chapter was funded by grants of the Deutsche Forschungsgemeinschaft (SFB 594 TP B05 and He2803/2–3).

## References

1. Frey, T.G., and Mannella, C.A. (2000). The internal structure of mitochondria. *Trends Biochem Sci* 25:319–324.
2. Frey, T.G., Renken, C.W., and Perkins, G.A. (2002). Insight into mitochondrial structure and function from electron tomography. *Biochim Biophys Acta* 1555:196–203.
3. Benz, R. (1994). Permeation of hydrophilic solutes through mitochondrial outer membranes: review on mitochondrial porins. *Biochim Biophys Acta* 1197:167–196.
4. Herrmann, J.M., and Hell, K. (2005). Chopped, trapped or tacked–protein translocation into the IMS of mitochondria. *Trends Biochem Sci* 30:205–212.
5. Koehler, C.M. (2004). New developments in mitochondrial assembly. *Annu Rev Cell Dev Biol* 20:309–335.
6. Stuart, R.A., Fölsch, H., Gruhler, A., and Neupert, W. (1996). Sorting of proteins to the mitochondrial intermembrane space, in Hartl, F.-U. (ed.). *Adv Molec Cell Biol* pp. 193–220, London: Jai-Press.
7. Michaelis, G., Esser, K., Tursun, B., Stohn, J.P., Hanson, S., and Pratje, E. (2005). Mitochondrial signal peptidases of yeast: the rhomboid peptidase Pcp1 and its substrate cytochrome *c* peroxidase. *Gene* 354:58–63.
8. Glick, B.S., Brandt, A., Cunningham, K., Muller, S., Hallberg, R.L., and Schatz, G. (1992). Cytochromes $c_1$ and $b_2$ are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. *Cell* 69:809–822.
9. Burri, L., Strahm, Y., Hawkins, C.J., Gentle, I.E., Puryer, M.A., Verhagen, A., Callus, B., Vaux, D., and Lithgow, T. (2005). Mature DIABLO/Smac is produced by the IMP protease complex on the mitochondrial inner membrane. *Mol Biol Cell* 16:2926–2933.
10. Nicholson, D.W., and Neupert, W. (1989). Import of cytochrome *c* into mitochondria: reduction of heme, mediated by NADH and flavin nucleotides, is obligatory for its covalent linkage to apocytochrome *c*. *Proc Natl Acad Sci USA* 86:4340–4344.
11. Mayer, A., Neupert, W., and Lill, R. (1995). Translocation of apocytochrome *c* across the outer membrane of mitochondria. *J Biol Chem* 270:12390–12397.
12. Dumont, M.E., Ernst, J.F., and Sherman, F. (1988). Coupling of heme attachment to import of cytochrome *c* into yeast mitochondria. *J Biol Chem* 263:15928–15937.
13. Nargang, F.E., Drygas, M.E., Kwong, P.L., Nicholson, D.W., and Neupert, W. (1988). A mutant of *Neurospora crassa* deficient in cytochrome *c* heme lyase activity cannot import cytochrome *c* into mitochondria. *J Biol Chem* 263:9388–9394.
14. Dumont, M.E., Cardillo, T.S., Hayes, M.K., and Sherman, F. (1991). Role of cytochrome *c* heme lyase in mitochondrial import and accumulation of cytochrome *c* in *Saccharomyces cerevisiae*. *Mol Cell Biol* 11:5487–5496.

15. Diekert, K., de Kroon, A.I., Ahting, U., Niggemeyer, B., Neupert, W., de Kruijff, B., and Lill, R. (2001). Apocytochrome *c* requires the TOM complex for translocation across the mitochondrial outer membrane. *EMBO J* 20:5626–5635.
16. Field, L.S., Furukawa, Y., O'Halloran, T.V., and Culotta, V.C. (2003). Factors controlling the uptake of yeast copper/zinc superoxide dismutase into mitochondria. *J Biol Chem* 278:28052–28059.
17. Steiner, H., Zollner, A., Haid, A., Neupert, W., and Lill, R. (1995). Biogenesis of mitochondrial heme lyases in yeast: import and folding in the intermembrane space. *J Biol Chem* 270:22842–22849.
18. Rojo, M., Hovius, R., Demel, R.A., Nicolay, K., and Wallimann, T. (1991). Mitochondrial creatine kinase mediates contact formation between mitochondrial membranes. *J Biol Chem* 266:20290–29295.
19. Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B., Rehling, P., Pfanner, N., *et al.* (2003). The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc Natl Acad Sci USA* 100:13207–13212.
20. Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J* 23:3735–3746.
21. Terziyska, N., Lutz, T., Kozany, C., Mokranjac, D., Mesecke, N., Neupert, W., Herrmann, J.M., and Hell, K. (2005). Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions. *FEBS Lett* 579:179–184.
22. Naoe, M., Ohwa, Y., Ishikawa, D., Ohshima, C., Nishikawa, S., Yamamoto, H., and Endo, T. (2004). Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *J Biol Chem* 279:47815–47821.
23. Hofmann, S., Rothbauer, U., Muhlenbein, N., Baiker, K., Hell, K., and Bauer, M.F. (2005). Functional and mutational characterization of human MIA40 acting during import into the mitochondrial intermembrane space. *J Mol Biol* 353:517–528.
24. Allen, S., Balabanidou, V., Sideris, D.P., Lisowsky, T., and Tokatlidis, K. (2005). Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome *c*. *J Mol Biol* 353:937–944.
25. Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J.M. (2005). A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121:1059–1069.
26. Lisowsky, T. (1992). Dual function of a new nuclear gene for oxidative phosphorylation and vegetative growth in yeast. *Mol Gen Genet* 232:58–64.
27. Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K.A., and Starzl, T.E. (1994). Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proc Natl Acad Sci USA* 91:8142–8146.
28. Coppock, D.L., and Thorpe, C. (2006). Multidomain flavin-dependent sulfhydryl oxidases. *Antioxid Redox Signal* 8:300–311.
29. Polimeno, L., Lisowsky, T., and Francavilla, A. (1999). From yeast to man–from mitochondria to liver regeneration: a new essential gene family. *Ital J Gastroenterol Hepatol* 31:494–500.
30. Lee, J., Hofhaus, G., and Lisowsky, T. (2000). Erv1p from *Saccharomyces cerevisiae* is a FAD-linked sulfhydryl oxidase. *FEBS Lett* 477:62–66.

31. Gerber, J., Muhlenhoff, U., Hofhaus, G., Lill, R., and Lisowsky, T. (2001). Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J Biol Chem* 276:23486–23491.
32. Hofhaus, G., Lee, J.E., Tews, I., Rosenberg, B., and Lisowsky, T. (2003). The N-terminal cysteine pair of yeast sulfhydryl oxidase Erv1p is essential for *in vivo* activity and interacts with the primary redox centre. *Eur J Biochem* 270:1528–1535.
33. Hoober, K.L., Sheasley, S.L., Gilbert, H.F., and Thorpe, C. (1999). Sulfhydryl oxidase from egg white. A facile catalyst for disulfide bond formation in proteins and peptides. *J Biol Chem* 274:22147–22150.
34. Sevier, C.S., Cuozzo, J.W., Vala, A., Aslund, F., and Kaiser, C.A. (2001). A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat Cell Biol* 3:874–882.
35. Hofhaus, G., Stein, G., Polimeno, L., Francavilla, A., and Lisowsky, T. (1999). Highly divergent amino termini of the homologous human ALR and yeast scERV1 gene products define species specific differences in cellular localization. *Eur J Cell Biol* 78:349–356.
36. Gross, E., Sevier, C.S., Vala, A., Kaiser, C.A., and Fass, D. (2002). A new FAD-binding fold and intersubunit disulfide shuttle in the thiol oxidase Erv2p. *Nat Struct Biol* 9:61–67.
37. Stein, G., and Lisowsky, T. (1998). Functional comparison of the yeast scERV1 and scERV2 genes. *Yeast* 14:171–180.
38. Vala, A., Sevier, C.S., and Kaiser, C.A. (2005). Structural determinants of substrate access to the disulfide oxidase Erv2p. *J Mol Biol* 354:952–966.
39. Wu, C.K., Dailey, T.A., Dailey, H.A., Wang, B.C., and Rose, J.P. (2003). The crystal structure of augmenter of liver regeneration: a mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci* 12:1109–1118.
40. Gross, E., Kastner, D.B., Kaiser, C.A., and Fass, D. (2004). Structure of Ero1p, source of disulfide bonds for oxidative protein folding in the cell. *Cell* 117:601–610.
41. Pawlowski, R., and Jura, J. (2006). ALR and liver regeneration. *Mol Cell Biochem* 288 (1–2):159–169.
42. Coppock, D.L., Cina-Poppe, D., and Gilleran, S. (1998). The quiescin Q6 gene (QSCN6) is a fusion of two ancient gene families: thioredoxin and ERV1. *Genomics* 54:460–468.
43. Senkevich, T.G., White, C.L., Koonin, E.V., and Moss, B. (2002). Complete pathway for protein disulfide bond formation encoded by poxviruses. *Proc Natl Acad Sci USA* 99:6667–6672.
44. Lewis, T., Zsak, L., Burrage, T.G., Lu, Z., Kutish, G.F., Neilan, J.G., and Rock, D.L. (2000). An African swine fever virus ERV1-ALR homologue, 9GL, affects virion maturation and viral growth in macrophages and viral virulence in swine. *J Virol* 74:1275–1285.
45. Senkevich, T.G., White, C.L., Koonin, E.V., and Moss, B. (2000). A viral member of the ERV1/ALR protein family participates in a cytoplasmic pathway of disulfide bond formation. *Proc Natl Acad Sci USA* 97:12068–12073.
46. Lange, H., Lisowsky, T., Gerber, J., Mühlenhoff, U., Kispal, G., and Lill, R. (2001). An essential function of the mitochondrial sulfhydryl oxidase Erv1p/ALR in the maturation of cytosolic Fe/S proteins. *EMBO Rep* 2:715–720.
47. Becher, D., Kricke, J., Stein, G., and Lisowsky, T. (1999). A mutant for the yeast scERV1 gene displays a new defect in mitochondrial morphology and distribution. *Yeast* 15:1171–1181.
48. Lisowsky, T. (1994). *ERV1* is involved in the cell-division cycle and the maintenance of mitochondrial genomes in *Saccharomyces cerevisiae*. *Curr Genet* 26:15–20.

49. Rissler, M., Wiedemann, N., Pfannschmidt, S., Gabriel, K., Guiard, B., Pfanner, N., and Chacinska, A. (2005). The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J Mol Biol* 353:485–492.
50. Lill, R., and Mühlenhoff, U. (2005). Iron-sulfur-protein biogenesis in eukaryotes. *Trends Biochem Sci* 30:133–141.
51. Dimmer, K.S., Fritz, S., Fuchs, F., Messerschmitt, M., Weinbach, N., Neupert, W., and Westermann, B. (2002). Genetic basis of mitochondrial function and morphology in *Saccharomyces cerevisiae*. *Mol Biol Cell* 13:847–853.
52. Tokatlidis, K. (2005). A disulfide relay system in mitochondria. *Cell* 121:965–967.
53. Lutz, T., Neupert, W., and Herrmann, J.M. (2003). Import of small Tim proteins into the mitochondrial intermembrane space. *EMBO J* 22:4400–4408.
54. Allen, S., Lu, H., Thornton, D., and Tokatlidis, K. (2003). Juxtaposition of the two distal $Cx_3C$ motifs via intrachain disulfide bonding is essential for the folding of Tim10. *J Biol Chem* 278:38505–38513.
55. Farrell, S.R., and Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome *c* reductase activity. *Biochemistry* 44:1532–1541.
56. Koehler, C.M. (2004). The small Tim proteins and the twin $Cx_3C$ motif. *Trends Biochem Sci* 29:1–4.
57. Bauer, M.F., Hofmann, S., Neupert, W., and Brunner, M. (2000). Protein translocation into mitochondria: the role of TIM complexes. *Trends Cell Biol* 100:25–31.
58. Hofmann, S., Rothbauer, U., Muhlenbein, N., Neupert, W., Gerbitz, K.D., Brunner, M., and Bauer, M.F. (2002). The C66W mutation in the deafness dystonia peptide 1 (DDP1) affects the formation of functional DDP1. TIM13 complexes in the mitochondrial intermembrane space. *J Biol Chem* 277:23287–23293.
59. Roesch, K., Curran, S.P., Tranebjaerg, L., and Koehler, C.M. (2002). Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a-TIMM13 complex. *Hum Mol Genet* 11:477–486.
60. Webb, C.T., Gorman, M.A., Lazarou, M., Ryan, M.T., and Gulbis, J.M. (2006). Crystal structure of the mitochondrial chaperone TIM9*10 reveals a six-bladed alpha-propeller. *Mol Cell* 21:123–133.
61. Lu, H., Allen, S., Wardleworth, L., Savory, P., and Tokatlidis, K. (2004). Functional TIM10 chaperone assembly is redox-regulated *in vivo*. *J Biol Chem* 279:18952–18958.
62. Curran, S.P., Leuenberger, D., Oppliger, W., and Koehler, C.M. (2002). The Tim9p-Tim10p complex binds to the transmembrane domains of the ADP/ATP carrier. *EMBO J* 21:942–953.
63. Curran, S.P., Leuenberger, D., Schmidt, E., and Koehler, C.M. (2002). The role of the Tim8p-Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *J Cell Biol* 158:1017–1027.
64. Sirrenberg, C., Endres, M., Fölsch, H., Stuart, R.A., Neupert, W., and Brunner, M. (1998). Carrier protein import into mitochondria mediated by the intermembrane proteins Tim10/Mrs11p and Tim12/Mrs5p. *Nature* 391:912–915.
65. Paschen, S.A., Rothbauer, U., Kaldi, K., Bauer, M.F., Neupert, W., and Brunner, M. (2000). The role of the TIM8–13 complex in the import of Tim23 into mitochondria. *EMBO J* 19:6392–6400.
66. Vergnolle, M.A., Baud, C., Golovanov, A.P., Alcock, F., Luciano, P., Lian, L.Y., and Tokatlidis, K. (2005). Distinct domains of small Tims involved in subunit interaction and substrate recognition. *J Mol Biol* 351:839–849.
67. Curran, S.P., Leuenberger, D., Leverich, E.P., Hwang, D.K., Beverly, K.N., and Koehler, C.M. (2004). The role of Hot13p and redox chemistry in the mitochondrial TIM22 import pathway. *J Biol Chem* 279:43744–43751.

68. Lu, H., and Woodburn, J. (2005). Zinc binding stabilizes mitochondrial Tim10 in a reduced and import-competent state kinetically. *J Mol Biol* 353:897–910.
69. Arnesano, F., Balatri, E., Banci, L., Bertini, I., and Winge, D.R. (2005). Folding studies of Cox17 reveal an important interplay of cysteine oxidation and copper binding. *Structure* 13:713–722.
70. Glerum, D.M., Shtanko, A., and Tzagoloff, A. (1996). Characterization of *COX17*, a yeast gene involved in copper metabolism and assembly of cytochrome oxidase. *J Biol Chem* 271:14504–14509.
71. Beers, J., Glerum, D.M., and Tzagoloff, A. (1997). Purification, characterization, and localization of yeast Cox17p, a mitochondrial copper shuttle. *J Biol Chem* 272:33191–33196.
72. Abajian, C., Yatsunyk, L.A., Ramirez, B.E., and Rosenzweig, A.C. (2004). Yeast cox17 solution structure and Copper(I) binding. *J Biol Chem* 279:53584–53592.
73. Srinivasan, C., Posewitz, M.C., George, G.N., and Winge, D.R. (1998). Characterization of the copper chaperone Cox17 of *Saccharomyces cerevisiae*. *Biochemistry* 37:7572–7577.
74. Lamb, A.L., Wernimont, A.K., Pufahl, R.A., Culotta, V.C., O'Halloran, T.V., and Rosenzweig, A.C. (1999). Crystal structure of the copper chaperone for superoxide dismutase. *Nat Struct Biol* 6:724–729.
75. Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Gonnelli, L., and Mangani, S. (2004). Solution structure of Cox11, a novel type of beta-immunoglobulin-like fold involved in CuB site formation of cytochrome *c* oxidase. *J Biol Chem* 279:34833–34839.
76. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., and Yoshikawa, S. (1995). Structures of metal sites of oxidized bovine heart cytochrome *c* oxidase at 2.8 A. *Science* 269:1069–1074.
77. Nobrega, M.P., Bandeira, S.C., Beers, J., and Tzagoloff, A. (2002). Characterization of *COX19*, a widely distributed gene required for expression of mitochondrial cytochrome oxidase. *J Biol Chem* 277:40206–40211.
78. Barros, M.H., Johnson, A., and Tzagoloff, A. (2004). *COX23*, a homologue of *COX17*, is required for cytochrome oxidase assembly. *J Biol Chem* 279:31943–31947.
79. Levitan, A., Danon, A., and Lisowsky, T. (2004). Unique features of plant mitochondrial sulfhydryl oxidase. *J Biol Chem* 279:20002–20008.
80. Iwata, S., Lee, J.W., Okada, K., Lee, J.K., Iwata, M., Rasmussen, B., Link, T.A., Ramaswamy, S., and Jap, B.K. (1998). Complete structure of the 11-subunit bovine mitochondrial cytochrome $bc_1$ complex. *Science* 281:64–71.
81. Chinenov, Y.V. (2000). Cytochrome *c* oxidase assembly factors with a thioredoxin fold are conserved among prokaryotes and eukaryotes. *J Mol Med* 78:239–242.
82. Lamb, A.L., Torres, A.S., O'Halloran, T.V., and Rosenzweig, A.C. (2001). Heterodimeric structure of superoxide dismutase in complex with its metallochaperone. *Nat Struct Biol* 8:751–755.

# 14

# *The Function of TIM22 in the Insertion of Inner Membrane Proteins in Mitochondria*

YAVUZ OKTAY • ROBERT N. RAINEY • CARLA M. KOEHLER

*Department of Chemistry and Biochemistry and the Molecular Biology Institute*
*University of California Los Angeles*
*Box 951569, Los Angeles, CA 90095, USA*

## I. Abstract

Diverse pathways accommodate the import of proteins into the mitochondrion. In contrast to the translocase of the inner membrane 23 (TIM23) import pathway utilized by matrix-targeted proteins, the TIM22 pathway mediates the import of polytopic inner membrane proteins such as the mitochondrial carrier family and import components Tim17p, Tim22p, and Tim23p. Substrates of the TIM22 pathway lack the typical N-terminal targeting sequence and instead contain targeting information within the mature part of the protein. Components of the TIM22 pathway include the small Tim proteins, Tim8p, Tim9p, Tim10p, Tim12p, and Tim13p, which are soluble components in the intermembrane space. Tim9p partners with Tim10p and Tim8p with Tim13p to form chaperone-like complexes and escort the precursor across the aqueous intermembrane space. At the inner membrane, the 300-kDa insertion complex consisting of Tim12p, Tim18p, Tim22p, and Tim54p and a fraction of the Tim9p and Tim10p mediates insertion into the inner membrane. The TIM22 pathway is evolutionarily conserved from yeast to plants and animals. However, homologues have not been identified in prokaryotes, suggesting that the TIM22 pathway developed after



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25014-0

endosymbiosis. Interestingly, the first mitochondrial disease (Mohr-Tranebjaerg or deafness-dystonia syndrome) associated with a defect in protein import is caused by loss-of-function mutations in one of the small Tim proteins, TIMM8A/DDP1. Identification of the TIM22 pathway confirms that mitochondrial protein import and assembly pathways are indeed complex.

## II. Introduction

The mitochondrion contains translocons on both the outer and inner membrane [1–4]. Mitochondrial precursors pass through the translocase of the outer membrane (TOM) complex and those destined for the intermembrane space, inner membrane, and matrix then take a variety of pathways. This chapter will focus on the biogenesis of polytopic inner membrane proteins that utilize the translocase of the inner membrane 22 (TIM22, the major component is the 22-kDa Tim22 protein) pathway, which represents a new addition to the import pathways in the mitochondrion [5–8]. In contrast, the TIM23 pathway was the first import pathway characterized, and it was assumed that the TIM23 complex accommodated all proteins except those of the outer membrane [9, 10]. However, mitochondrial biogenesis is indeed complex and the TIM22 represents one of the latest discoveries.

The mitochondrion contains a genome that codes for a small number of proteins that assemble in the respiratory complexes in the inner membrane [11]. During endosymbiosis, most genes moved from the ''progenitor mitochondrion'' to the nucleus. Subsequently, these proteins had to develop targeting sequences to return to the mitochondrion and the organelle had to develop translocons [11, 12]. The outer membrane contains the TOM complex. Generally all proteins that are destined for the mitochondrion utilize the TOM complex [9, 10]. Most mitochondrial precursors contain a typical N-terminal targeting sequence that directs them to the TIM23 translocon in the inner membrane [13, 14]. Here translocation is dependent on the presence of a membrane potential. However, many inner membrane proteins lack a typical N-terminal targeting sequence and instead carry targeting information within the mature part of the protein; these precursors utilize the TIM22 pathway for biogenesis [5–8].

The TIM22 pathway differs from the TIM23 pathway in that components of the TIM22 pathway are both soluble in the intermembrane space and embedded in the inner membrane (Figure 14.1; Table 14.1). In the intermembrane space, the small Tim family acts as chaperone-like complexes to guide precursors across the intermembrane space [6, 15]. The complement of small Tim proteins is 5 in yeast and 6 in mouse and human. At the inner membrane, the 300-kDa membrane complex consists of the pore-forming

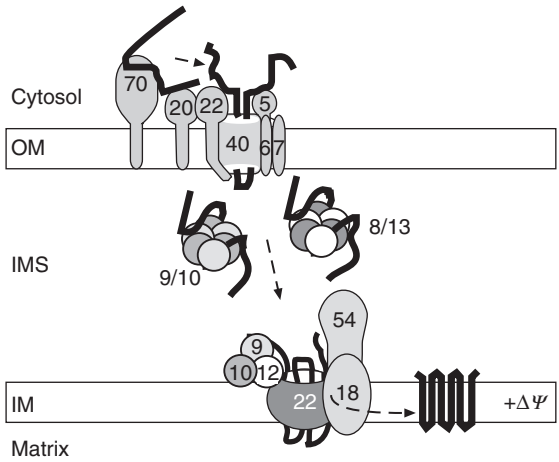



FIG. 14.1. The TIM22 translocation pathway. The TIM22 translocon mediates the import of inner membrane proteins such as those of the mitochondrial carrier family, Tim22p, and Tim23p. The small Tim proteins form 70-kDa complexes in the intermembrane space; Tim9p partners with Tim10p and Tim8p partners with Tim13p. The 300-kDa complex at the inner membrane contains intermembrane space protein Tim12p with a fraction of the Tim9p and Tim10p and the inner membrane components Tim18p, Tim22p, and Tim54p. Tim22p forms the translocation pore while Tim18p and Tim54p perform accessory functions.

TABLE 14.1

COMPONENTS OF THE TIM22 TRANSLOCON

| Component | Essential for Viability | Membrane Association | General Function* | References |
|---|---|---|---|---|
| Tim54p | No* | Integral | Translocon assembly? | [83] |
| Tim22p | Yes | Integral | Translocon pore | [27] |
| Tim18p | No | Integral | Translocon assembly? | [84, 85] |
| Tim13p | No | Soluble | Chaperone | [25, 59] |
| Tim12p | Yes | Peripheral | Chaperone | [63, 80] |
| Tim10p | Yes | Soluble | Chaperone | [63, 80] |
| Tim9p | Yes | Soluble | Chaperone | [58, 60] |
| Tim8p | No | Soluble | Chaperone | [25, 59] |

* indicates that the gene can only be deleted under certain conditions.
? indicates most likely function.

unit Tim22p, a fraction of the small Tim proteins (Tim9p, Tim10p, and Tim12), and accessory proteins Tim18p and Tim54p [5–8]. Insertion depends on the presence of a membrane potential, but energetic requirements for ATP seemingly are not required. In addition, matrix-sided components such as the Hsp70-driven translocation motor do not associate with the TIM22 machinery.

## III. Properties of Precursors that Utilize the TIM22 Import Pathway

The mitochondrial carrier family is a large protein family in the mitochondrial inner membrane that functions as metabolite transporters to mediate the passage of ions, nucleotides, and metabolites between the matrix and intermembrane space [16–18]. Yeast has ~35 carriers, and examples include the ADP/ATP carrier (AAC), the phosphate carrier (PiC), and the mammalian uncoupling proteins (UCPs). All carriers consist of 3 bipartite repeats of ~100 amino acids in which 2 transmembrane domains are separated by a positively charged loop (Figure 14.2A) [19]. As a result, the carriers contain six membrane-spanning domains and the N- and C-termini face the intermembrane space [20]. The positively charged regions facilitate the insertion of the carriers into the inner membrane, taking advantage of the electrophoretic effect of the inner membrane [21]; the intermembrane space side of the inner membrane is more positively charged than the negatively charged matrix side because of proton pumping by the respiratory complexes [22]. Biochemical studies show that the carriers function as dimers in the inner membrane [23]. After many years of biochemical analysis, a crystal structure for AAC in the monomeric form has been solved [20]. The crystal structure confirms the predicted topology in the inner membrane from biochemical studies.

In addition, the import components Tim22p, Tim23p, and most likely Tim17p utilize the TIM22 pathway for import (Figure 14.2A) [24–26]. In contrast to the carrier proteins, Tim22p, Tim23p, and Tim17p contain four predicted membrane-spanning domains and the N- and C-termini face the intermembrane space [24, 27]. Tim23p has an additional 10-kDa domain on the N-terminus that folds into a soluble domain [28]. Tim22p, Tim23p, and Tim17p share an identity of 15–25% and show low similarity to the LivH permease of the inner membrane of proteobacteria and the chloroplast outer envelope import component Oep16 [29, 30]. The proteobacterial protein transports branched chain amino acids. Therefore, Tim17p, Tim22p, Tim23p, Oep16, and LivH are all members of the preprotein and amino acid transporter motif (PRAT) family and may have evolved from a common eubacteria [12].

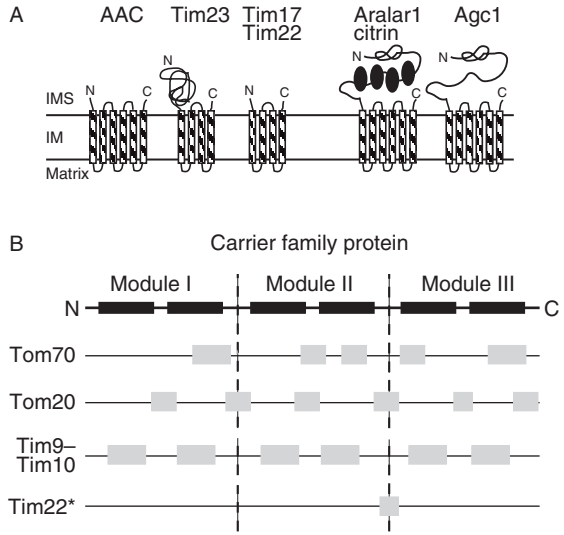



FIG. 14.2. Topology of inner membrane substrates of the TIM22 pathway and proposed regions of substrate-import component interaction. (A) Topological model of substrates of the TIM22 pathway. Proteins depicted include the AAC; import components Tim17p, Tim22p, and Tim23p, and AGCs mammalian Aralar1 and citrin and yeast Agc1p. The black circles on Aralar1 and citrin represent calcium-binding EF hands. (B) A linear representation of a typical carrier protein showing the modules that contain two transmembrane domains. The gray boxes represent binding regions for Tom70p [33], Tom20p [33], and Tim9p–Tim10p [34–36, 38] as deduced by peptide scans. The specific binding sites for Tim22p have not been defined, but a peptide that binds to Tim22p has been marked.

Where is the targeting information in the substrates? The carrier proteins lack the typical N-terminal targeting sequence; however, a few carriers oddly do contain an extension at the N-terminus, but these sequences have been shown to be dispensable for import and assembly [31, 32]. The targeting information is located throughout the carrier proteins as shown by peptide scan experiments (Figure 14.2B) [33–36]. In peptide scans, sequential peptides of ~13 amino acids with a 10-amino acid overlap are spotted on a membrane [37]. The protein of interest, such as the Tim9p–Tim10p complex, is incubated with the peptide scan, and after washing, protein binding is detected by immunoblot analysis and quantitated. By using this approach, distinct binding sites for the TOM complex and the small Tim complexes were found throughout the carrier proteins [33–36]. Specifically, Tom70p bound to hydrophobic regions of the carrier, whereas Tom20p bound to hydrophilic regions [33]. Tom70p and Tom20p are two receptors at the outer membrane that direct the precursor to the TOM channel. The small Tim proteins showed strongest binding in the membrane-spanning

domains [34–36, 38], suggesting that the small Tim proteins maintain the hydrophobic proteins in an import-competent state in the aqueous intermembrane space.

Whereas binding sites in the carrier proteins for the TOM receptors and small Tim proteins have been mapped, interactions with the 300-kDa membrane complex have not been specifically identified. The carriers have been dissected into modules for *in organello* import assays [39, 40]; the targeting information seems to reside in Module III (Figure 14.2B). However, because these experiments use truncated forms of the carriers, their association with the translocons might be altered. Similarly, peptide scan experiments in which integral membrane proteins such as Tim22p are used as the probes are difficult because of folding and solubility problems. An uncharged peptide from the intermembrane space loop between Modules II and III binds to reconstituted Tim22p and affects channel activity [41]. Thus, specific regions of substrates can be expected to interact with Tim22p, but have not been rigorously identified.

In general, precursors with an N-terminal targeting sequence cross the TOM complex as an extended chain with the N-terminus entering the TOM channel first. How do the TIM22 substrates negotiate the TOM channel? Studies in which folded domains were appended to the N- and C-termini of AAC or Tim23p demonstrate that the substrates crossed the TOM complex as a loop [34, 35, 42]. The TOM channel stretches to 26 Å [43, 44] so it can accommodate two unfolded polypeptide chains in the translocon at one time. Thus, the TIM22 precursors negotiate the TOM complex differently than precursors with an N-terminal targeting sequence, which may serve as a sorting mechanism to divert a substrate to the correct translocon.

The inner membrane also contains inner membrane proteins in which the biogenesis pathway has not been elucidated. As an example, the ABC transporters Atm1p, Mdl1p, and Mdl2p are predicted to share a similar topology in the inner membrane of six membrane-spanning domains with the N- and C-termini facing the matrix [45]. Atm1p and Mdl1p contain typical N-terminal targeting sequences, whereas the N-terminus of Mdl2p is not predicted to direct mitochondrial localization [45]. The import of subunit e of the ATPase/Tim11p [46, 47], with one transmembrane domain, may be translocated spontaneously or via a new mechanism because import seemed independent of the TIM22 and TIM23 translocons [48]. Finally, a new class of membrane protein is represented by the putative cardiolipin remodeling enzyme, Taz1p, which contains one hydrophobic domain [49, 50]. Topology and localization studies indicate that Taz1p lines the intermembrane space in the outer and inner membrane [49]. Rather than traversing the membrane, the hydrophobic domain anchors the protein to the membrane with the N- and C-termini facing the intermembrane space [49].

Biogenesis studies suggest that Taz1p may be imported to the outer membrane, relying on the small Tim proteins [50]. Thus, many details about the mechanism of membrane protein translocation and insertion remain to be elucidated.

## IV. The Small Tim Proteins

The small Tim proteins or the ''tiny Tims'' are a family of proteins in the intermembrane space that share an identity of ~25%, similarity of 40%, and molecular weight near 10 kDa [6, 51]. The proteins are not similar to other known proteins but contain an interesting ''twin CX3C'' motif in which two cysteine residues are separated by three amino acids [52, 53]. Spacing between each CX3C motif varies from 11 to 16 amino acids. Yeast have five small Tim proteins (Tim8p, Tim9p, Tim10p, Tim12p, and Tim13p), whereas humans possess six (DDP1/TIMM8A, DDP2/TIMM8B, TIMM9A, TIMM9B, TIMM10, and TIMM13) [6, 54]. The cysteine residues are important for function because a mutation in one cysteine in DDP1 results in deafness-dystonia syndrome [55, 56]. In yeast, Tim9p, Tim10p, and Tim12p are essential for viability, whereas deletion of TIM8 and TIM13 does not significantly impair growth under normal conditions. However, in certain strain backgrounds, loss of *TIM8* and *TIM13* display cold sensitivity [25, 57].

The interactions between the small Tim proteins are complex [51 58–60]. Tim9p and Tim10p have two locations within the intermembrane space. Approximately 95% of the Tim9p and Tim10p pair to form a 70-kDa complex that is soluble in the intermembrane space, and the remainder associates with the TIM22 membrane complex [61]. Likewise, Tim8p partners with Tim13p to form a similar soluble complex [25, 57, 59]. In contrast, Tim12p remains associated with the TIM22 membrane complex. The subunit ratio in the 70-kDa complexes is 3:3. The twin-CX3C motif seems like it may coordinate a metal. Initial studies showed that the recombinant monomeric small Tim proteins indeed bound zinc, forming a zinc fingerlike structure [57, 62, 63]. Likewise, import of the small Tims into the intermembrane space requires the four conserved thiol groups and zinc ions are potentially critical for stable folding and assembly [64]. However, the spacing of three residues between the cysteines differs from the canonical zinc finger that typically has a spacing of two residues [65], suggesting that the motif might have alternative properties.

In contrast to the metal-binding reports, a set of different studies demonstrated that the recombinant and native Tim9p–Tim10p and Tim8p–Tim13p complexes did not coordinate zinc, but instead formed disulfide

linkages [34, 35, 53]. Specifically, the cysteine residues paired in juxtaposed disulfide bonds [53]. Whereas the fully reduced Tim proteins can bind zinc, binding does not promote complex formation in reconstitution experiments with *Neurospora crassa* TOM complex in lipid vesicles; the trapped Tim9p–Tim10p complex forms disulfide linkages and promotes translocation of AAC [38]. Zinc coordination may be required for the import and assembly of the small Tim proteins in the intermembrane space in yet another pathway consisting of intermembrane space proteins Erv1p and Mia40p [66–71]. Definitively, the first structure of the Tim9p–Tim10p complex has been determined and confirms that the cysteine residues form intramolecular disulfide bonds [72]. The structure unexpectedly resembles that of the Skp/Prefoldin chaperones of the bacterial periplasm, although there is no sequence similarity [72]. As the structures of additional mitochondrial components are determined, a better picture of the mechanistic interactions will become evident.

The Tim9p–Tim10p and Tim8p–Tim13p complexes have different substrate specificities. The Tim9p–Tim10p complex predominantly binds to carrier proteins and import components Tim17p and Tim22p [48, 58]. In contrast, the Tim8p–Tim13p complex binds to Tim23p, but Tim9p and Tim10p can also be cross-linked to an arrested Tim23p import precursor [24, 48, 57]. The small Tim complexes bind to hydrophobic regions of the substrates [34, 35, 38]. In addition, the reconstituted Tim9p–Tim10p complex can prevent aggregation and mediate refolding of a model substrate [73]. Obvious substrate-binding pockets, however, were not detected in the structural analysis of the small Tim complex [72], suggesting that the complex may undergo conformational changes when a substrate binds. Hence, the small Tim complexes act as chaperones to prevent the hydrophobic substrates from aggregating in the aqueous intermembrane space.

As the inner membrane substrates exit the TOM channel in an unfolded conformation, a loop is exposed to the aqueous intermembrane space [34, 35]. The small Tim complexes then bind the precursor to escort it to the Tim22p insertion complex in the inner membrane. In the carrier proteins, the last transmembrane domain is important for translocation across the TOM complex [39, 40]. Just as a direct interaction between the TOM and TIM23 complex was not observed [74], direct binding between the TOM complex and small Tims has not been observed by standard biochemical methods [1].

Additional studies suggest that the small Tim complexes act as a general chaperone complex in the intermembrane space to assist the import of hydrophobic proteins. Cross-linking analysis illustrated that the small Tim proteins bind directly to a Tom40p translocation intermediate that is exposed to the intermembrane space during transport from the TOM

complex to the SAM complex [75, 76]. The SAM complex mediates insertion of β-barrel proteins into the outer membrane [77, 78]; the β-barrel proteins including porin, Tom40p, and Sam50p are trafficked from the TOM complex to the SAM complex via the intermembrane space. Whereas the small Tim proteins facilitate import of the β-barrel substrates, loss of function mutations in the small Tim proteins do not impair assembly of outer membrane complexes [75]. In addition, interactions between the prohibitins and Tim8p–Tim13p have been trapped [79]; the prohibitins are anchored to the inner membrane, facing the intermembrane space. However, from genetic approaches with yeast mutants, assembly of the carrier family and Tim22p and Tim23p proteins seems to be most severely affected when the function of small Tim proteins is impaired [59, 60, 80]. Thus the small Tim proteins act as general chaperones in the intermembrane space, but seem to play a specific role in assembly of the carrier proteins, Tim22p and Tim23p, which may be reflected by their specific association with the 300-kDa Tim22p complex at the inner membrane [61].

## V. The TIM22 Inner Membrane Complex

The inner membrane substrates are guided to the 300-kDa insertion complex at the inner membrane that contains Tim12p, Tim18p, Tim22p, and Tim54p and a fraction of Tim9p and Tim10p [6, 81]. Although Tim22p shares sequence similarity with Tim23p and Tim17, these translocons cannot substitute for each other [27]. Tim12p and Tim22p are essential for viability [27, 82]; Tim54p also was deemed essential for viability [83], but additional studies have shown that it can be deleted under certain conditions resulting in severely compromised growth [41]. In contrast, Tim18p is not essential [84, 85], but strains lacking Tim18p exhibit cold sensitivity and petite negativity [85, 86]. Tim22p mediates the insertion of carriers into the inner membrane in the presence of a membrane potential; a translocation intermediate has been arrested in association with Tim22p, presumably in the translocation channel, when the membrane potential was lowered by treatment with uncoupling agents [39, 87].

Pfanner and colleagues [87] have analyzed the TIM22 translocation pore; a ''minimal'' Tim22p translocon has been purified from a yeast strain that is viable without Tim54p. Biochemical characterization of the reconstituted Tim22 translocon and recombinant Tim22p in liposomes has shown that Tim22p forms a channel with multiple conductance states and more than one pore [87]. In its most open state of 18 Å, the Tim22p pore could facilitate the insertion of two tightly packed α-helices, whereas the intermediate confirmation of 11 Å could accommodate one single

transmembrane helix. How these two open states are regulated to facilitate translocation has not been elucidated; however, the current model would predict that in the most open state a loop could be inserted across the inner membrane and, following lateral diffusion of one transmembrane domain from the translocation pore, the channel would constrict to the smaller conformation to maintain the membrane potential across the inner membrane [87]. Until structural studies or additional molecular experiments are presented, the specific mechanism remains to be elucidated.

Less is known about the function of accessory proteins Tim54p and Tim18p. In addition, these accessory proteins do not seem to be conserved in higher eukaryotes, suggesting a fungal-specific function [12]. Interestingly, Tim54p was identified in a two-hybrid screen with the morphology component Mmm1p, but additional biochemical analysis showed that Tim22p and Tim54p instead are partner proteins [83]. Tim54p is anchored to the inner membrane by a hydrophobic domain at the N-terminus and most of the protein folds into a domain in the intermembrane space. A conditional allele of *tim54* abrogated the import of AAC, but a direct binding interaction was not detected [83], suggesting that Tim54p may not directly facilitate import of precursors. Whereas Tim54p was initially believed to perform an essential role in protein import, Pfanner, Jensen and colleagues [41] have shown that deletion of *tim54* yields a viable, albeit extremely sick yeast strain. Moreover, loss of Tim54p did not compromise assembly of the core TIM22 translocon [41]. Tim54p, therefore, plays a peripheral role in protein import, perhaps mediating assembly of the complex. Alternatively, given that fungal metabolism is generally more diverse than metazoans, Tim54p may play a role under specific metabolic or stress conditions.

Tim18p is another nonessential component of the TIM22 translocon. First identified as a multicopy suppressor of a *tim54* conditional allele and as a binding partner with Tim54p, Tim18p is predicted to span the inner membrane three times and is targeted to the mitochondrion by a classical N-terminal presequence [84, 85]. Tim18p may play a role in assembly because loss of Tim18p results in a smaller TIM22 translocon of 250 kDa [84, 85]. In addition, mutants lacking Tim18p are inviable when the mitochondrial genome is lost (referred to as petite negative) and display a cold-sensitive phenotype on rich glucose media [85, 86]. Studies suggest that Tim18p might be involved in a genetic pathway from the cytosol to the mitochondrial matrix for maintaining viability when the mitochondrial genome is lost [86]; the proposed pathway increases mitochondrial translocation under conditions of mitochondrial stress. Additional biochemical and genetic studies will inevitably assign more specific roles for Tim18p and Tim54p in mitochondrial biogenesis.

## VI. Disease Connections

The first inherited disease caused by a defect in protein import is associated with a mutation in the TIM22 import pathway [59]. Mohr-Tranebjaerg syndrome (MTS) or deafness-dystonia syndrome is an X-linked disease associated with deafness, blindness, and dystonia [88, 89]. When the disease was first characterized in the 1960s, sensorineural deafness was the prominent symptom but later investigations in the 1990s revealed that families had a wide array of neurological defects including dystonia, blindness, dysphagia, and mental deterioration [90]. Interestingly, this disease differed from typical mitochondrial diseases that affect both muscular (myopathy) and neural (neuropathy) tissues [91]. Also, the symptoms varied dramatically within and across families [90], making diagnosis difficult. Although this disease is considered rare, it confirms that mitochondrial biogenesis is important in the neural system. Additional diseases associated with defects in mitochondrial biogenesis may be difficult to identify because they are lethal. Finally, deciphering symptoms that are caused by a defect in protein import prove difficult because of the variability of symptoms. As methods are developed to more precisely map disease genes, additional mitochondrial diseases will most certainly be identified [92, 93].

MTS is caused by loss-of-function mutations in the small Tim protein, TIMM8A/DDP1 [89]. *DDP1* is on the X-chromosome near the *BTK1* gene. The locus was first identified by linkage analysis in a large family that contained a 26-kb deletion on the X-chromosome [89]; the C-terminus of the *BTK1* and the entire *DDP1* coding region were deleted. Subsequently, additional families with MTS have been characterized and most cases result in a loss-of-function mutation [94–97]. Interestingly, one family has a mutation in which the fourth cysteine of the twin-CX3C motif is mutated to a tryptophan (C66W) [56, 98]. This mutation was characterized in yeast and patient cell lines and the DDP1 protein and the DDP1-TIMM13 complex were not detected [55, 56].

A specific metabolic defect in target tissues might be the underlying cause of MTS. Expression analysis shows that DDP1 and TIMM13 are highly expressed in brain and liver and expressed to a lower extent in heart and muscle [99]. Specifically, prominent expression was detected in the soma and the dendritic portion of the Purkinje cells of the cerebellum, but not in the glial cells. Scattered expression also was detected in the brain stem, olfactory bulb, substantia nigra, hippocampus, and striatum [99]. From patient studies, defects in basal ganglia function contribute to symptoms of MTS [90, 95, 96, 100]. Specifically, pathological changes in the basal ganglia and sensory cortex demonstrated the disintegration of subcortico-cortical

circuits [95]. The small Tim proteins seem to display specific tissue expression patterns so different combinations of the small Tim proteins may be required for different tissues. From molecular studies in yeast mitochondria, the small Tim complexes have different substrate specificities [55, 99]. Additionally, the aspartate/glutamate carriers (AGCs), citrin and Aralar1, display similar expression patterns with DDP1 and TIMM13 [101–104]. From biochemical studies, DDP1 and TIMM13 along with Tim9p and Tim10p mediated the import of citrin and Aralar1 into mitochondria and were specifically cross-linked in an *in organello* import assay. Therefore, defects in the import of a specific subset of inner membrane substrates most likely contribute to the pathophysiology of MTS.

The AGCs are an interesting branch of the mitochondrial carrier family possessing an N-terminal extension. Whereas citrin and Aralar1 contain $Ca^{2+}$-responsive EF-binding hands in the N-terminal domain, the yeast Agc1p lacks the calcium-binding motif (Figure 14.2A) [101,105–107]. The AGCs are important in the Nicotinamide adenine dinucleotide (NADH) shuttle for moving reducing equivalents from the cytosol to the mitochondrion, potentially responding to calcium waves and allowing for rapid adjustments to metabolism [102, 104, 106]. Studies to define the underlying metabolic defect associated with MTS in patient fibroblasts have not been very successful because obvious energetic differences have not been identified between patient and control cell lines [55]. This may be reflected by the fact that fibroblast cell lines are not metabolically robust [108]. In addition, obvious differences were not detected in the assembly of the respiratory complexes in fibroblasts and muscles [55, 95]. In a lymphoblast cell line derived from the C66W MTS patient, the abundance of NADH and the AGCs was decreased [99]. This implies that a specific defect associated with NADH shuttling may contribute to MTS. Ultimately, until a suitable model is developed, the molecular basis for MTS in the neural system can only be speculated.

Additional defects in mitochondrial biogenesis have been recently linked to disease. Mutations in a homologue of Mdj2p and Tim14p/Pam18p result in the novel autosomal recessive condition, dilated cardiomyopathy with ataxia syndrome [109]. This disease, identified in the Canadian Dariusleut Hutterite population, is characterized by early onset dilated cardiomyopathy, ataxia, growth failure, and 3-methylglutaconic aciduria. Tim14p/Pam18p functions in the protein-associated translocation motor of the TIM23 translocon, implicating a defect in import via the TIM23 translocon. Recently, Alzheimer's disease (AD) and Parkinson's disease have been linked to the mitochondrion. Specifically, the putative mitochondrial kinase, PINK1, has been found mutated in an inherited form of parkinsonism. Mutations in PINK1 confer different autophosphorylation activity, which is regulated by the C-terminal portion of the protein [110].

Potentially altered association or activity of PINK1 with mitochondria may impact the development of parkinsonism. Interestingly, nonglycosylated full-length and C-terminal-truncated amyloid precursor protein (APP) has been found to accumulate exclusively in the protein import channels of mitochondria of human AD brains but not in age-matched controls [111]. Furthermore, the mitochondrial-associated APP formed stable complexes of 480-kDa complexes with the TOM complex and 620 kDa with the TOM-TIM23 supercomplex. As expected, accumulated APP inhibited the import of other mitochondrial precursors, leading to impaired mitochondrial function, specifically complex I, and subsequent cellular stress. While this study reports the first attenuation of protein import leading to a disease, additional studies should provide clues into the role of protein translocation in the progression of AD [111].

The characterization of protein translocation in mammalian mitochondria is a new frontier in mitochondrial biology. How defects in protein import are linked to disease is just beginning to be understood. As improvements in gene linkage analysis and new high-throughput technologies develop with studies in model organisms such as worm, mouse, fly, and zebrafish, mitochondrial dysfunction associated with defects in import should represent a new and exciting avenue of research.

## References

1. Koehler, C.M. (2004). New developments in mitochondrial assembly. *Annu Rev Cell Dev Biol* 20:309–335.
2. Wiedemann, N., Frazier, A.E., and Pfanner, N. (2004). The protein import machinery of mitochondria. *J Biol Chem* 279:14473–14476.
3. Mokranjac, D., and Neupert, W. (2005). Protein import into mitochondria. *Biochem Soc Trans* 33:1019–1023.
4. Endo, T., Yamamoto, H., and Esaki, M. (2003). Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles. *J Cell Sci* 116:3259–3267.
5. Kaldi, K., and Neupert, W. (1998). Protein translocation into mitochondria. *Biofactors* 8:221–224.
6. Koehler, C.M., Merchant, S., and Schatz, G. (1999). How membrane proteins travel across the mitochondrial intermembrane space. *Trends Biochem Sci* 24:428–432.
7. Pfanner, N. (1998). Mitochondrial import: crossing the aqueous intermembrane space. *Curr Biol* 8:R262–R265.
8. Truscott, K.N., Brandner, K., and Pfanner, N. (2003). Mechanisms of protein import into mitochondria. *Curr Biol* 13:R326–R337.
9. Pfanner, N., Craig, E.A., and Honlinger, A. (1997). Mitochondrial preprotein translocase. *Annu Rev Cell Dev Biol* 13:25–51.
10. Neupert, W. (1997). Protein import into mitochondria. *Annu Rev Biochem* 66:863–917.
11. Dolezal, P., Likic, V., Tachezy, J., and Lithgow, T. (2006). Evolution of the molecular machines for protein import into mitochondria. *Science* 313:314–318.

12. Lister, R., Hulett, J.M., Lithgow, T., and Whelan, J. (2005). Protein import into mitochondria: origins and functions today. *Mol Membr Biol* 22:87–100.
13. Neupert, W., and Brunner, M. (2002). The protein import motor of mitochondria. *Nat Rev Mol Cell Biol* 3:555–565.
14. Rehling, P., Wiedemann, N., Pfanner, N., and Truscott, K.N. (2001). The mitochondrial import machinery for preproteins. *Crit Rev Biochem Mol Biol* 36:291–336.
15. Bauer, M.F., Rothbauer, U., Muhlenbein, N., Smith, R.J., Gerbitz, K., Neupert, W., Brunner, M., and Hofmann, S. (1999). The mitochondrial TIM22 preprotein translocase is highly conserved throughout the eukaryotic kingdom. *FEBS Lett* 464:41–47.
16. Palmieri, F., Agrimi, G., Blanco, E., Castegna, A., Di Noia, M.A., Iacobazzi, V., Lasorsa, F.M., Marobbio, C.M., Palmieri, L., Scarcia, P., Todisco, S., Vozza, A., *et al.* (2006). Identification of mitochondrial carriers in *Saccharomyces cerevisiae* by transport assay of reconstituted recombinant proteins. *Biochim Biophys Acta* 1757:1249–1262.
17. Picault, N., Hodges, M., Palmieri, L., and Palmieri, F. (2004). The growing family of mitochondrial carriers in Arabidopsis. *Trends Plant Sci* 9:138–146.
18. Palmieri, F. (2004). The mitochondrial transporter family (SLC25): physiological and pathological implications. *Pflugers Arch* 447:689–709.
19. Palmieri, F., Bisaccia, F., Capobianco, L., Dolce, V., Fiermonte, G., Iacobazzi, V., Indiveri, C., and Palmieri, L. (1996). Mitochondrial metabolite transporters. *Biochim Biophys Acta* 1275:127–132.
20. Kunji, E.R., and Harding, M. (2003). Projection structure of the atractyloside-inhibited mitochondrial ADP/ATP carrier of *Saccharomyces cerevisiae*. *J Biol Chem* 278:36985–36988.
21. El Moualij, B., Duyckaerts, C., Lamotte-Brasseur, J., and Sluse, F.E. (1997). Phylogenetic classification of the mitochondrial carrier family of *Saccharomyces cerevisiae*. *Yeast* 13:573–581.
22. Sipos, L., and von Heijne, G. (1993). Predicting the topology of eukaryotic membrane proteins. *Eur J Biochem* 213:1333–1340.
23. Trezeguet, V., Le Saux, A., David, C., Gourdet, C., Fiore, C., Dianoux, A., Brandolin, G., and Lauquin, G.J. (2000). A covalent tandem dimer of the mitochondrial ADP/ATP carrier is functional *in vivo*. *Biochim Biophys Acta* 1457:81–93.
24. Davis, A.J., Ryan, K.R., and Jensen, R.E. (1998). Tim23p contains separate and distinct signals for targeting to mitochondria and insertion into the inner membrane. *Mol Biol Cell* 9:2577–2593.
25. Davis, A.J., Sepuri, N.B., Holder, J., Johnson, A.E., and Jensen, R.E. (2000). Two intermembrane space TIM complexes interact with different domains of Tim23p during its import into mitochondria. *J Cell Biol* 150:1271–1282.
26. Leuenberger, D., Curran, S.P., Wong, D., and Koehler, C.M. (2003). The role of Tim9p in the assembly of the TIM22 import complexes. *Traffic* 4:144–152.
27. Sirrenberg, C., Bauer, M.F., Guiard, B., Neupert, W., and Brunner, M. (1996). Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22. *Nature* 384:582–585.
28. Emtage, J.L., and Jensen, R.E. (1993). MAS6 encodes an essential inner membrane component of the yeast mitochondrial protein import pathway. *J Cell Biol* 122:1003–1012.
29. Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M., and Soll, J. (1999). The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J Mol Biol* 286:105–120.
30. Dyall, S.D., Brown, M.T., and Johnson, P.J. (2004). Ancient invasions: from endosymbionts to organelles. *Science* 304:253–257.

31. Zara, V., Palmieri, F., Mahlke, K., and Pfanner, N. (1992). The cleavable presequence is not essential for import and assembly of the phosphate carrier of mammalian mitochondria but enhances the specificity and efficiency of import. *J Biol Chem* 267:12077–12081.
32. Zara, V., Ferramosca, A., Papatheodorou, P., Palmieri, F., and Rassow, J. (2005). Import of rat mitochondrial citrate carrier (CIC) at increasing salt concentrations promotes presequence binding to import receptor Tom20 and inhibits membrane translocation. *J Cell Sci* 118:3985–3995.
33. Brix, J., Rudiger, S., Bukau, B., Schneider-Mergener, J., and Pfanner, N. (1999). Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein. *J Biol Chem* 274:16522–16530.
34. Curran, S.P., Leuenberger, D., Oppliger, W., and Koehler, C.M. (2002). The Tim9p-Tim10p complex binds to the transmembrane domains of the ADP-ATP carrier. *EMBO J* 21:942–953.
35. Curran, S.P., Leuenberger, D., Schmidt, E., and Koehler, C.M. (2002). The role of the Tim8p-Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. *J Cell Biol* 158:1017–1027.
36. Vergnolle, M.A., Sawney, H., Junne, T., Dolfini, L., and Tokatlidis, K. (2005). A cryptic matrix targeting signal of the yeast ADP/ATP carrier normally inserted by the TIM22 complex is recognized by the TIM23 machinery. *Biochem J* 385:173–180.
37. Reineke, U., Sabat, R., Kramer, A., Stigler, R.D., Seifert, M., Michel, T., Volk, H.D., and Schneider-Mergener, J. (1996). Mapping protein–protein contact sites using cellulose-bound peptide scans. *Mol Divers* 1:141–148.
38. Vasiljev, A., Ahting, U., Nargang, F.E., Go, N.E., Habib, S.J., Kozany, C., Panneels, V., Sinning, I., Prokisch, H., Neupert, W., Nussberger, S., and Rapaport, D. (2004). Reconstituted TOM core complex and Tim9/Tim10 complex of mitochondria are sufficient for translocation of the ADP/ATP carrier across membranes. *Mol Biol Cell* 15:1445–1458.
39. Brandner, K., Rehling, P., and Truscott, K.N. (2005). The carboxyl-terminal third of the dicarboxylate carrier is crucial for productive association with the inner membrane twin-pore translocase. *J Biol Chem* 280:6215–6221.
40. Endres, M., Neupert, W., and Brunner, M. (1999). Transport of the ADP/ATP carrier of mitochondria from the TOM complex to the TIM22. 54 complex. *EMBO J* 18:3214–3221.
41. Kovermann, P., Truscott, K.N., Guiard, B., Rehling, P., Sepuri, N.B., Muller, H., Jensen, R.E., Wagner, R., and Pfanner, N. (2002). Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. *Mol Cell* 9:363–373.
42. Ryan, M.T., Müller, H., and Pfanner, N. (1999). Functional Staging of ADP/ATP carrier translocation across the outer mitochondrial membrane. *J Biol Chem* 274:20619–20627.
43. Schwartz, M.P., and Matouschek, A. (1999). The dimensions of the protein import channel in the outer and inner mitochondrial membranes. *Proc Natl Acad Sci USA* 98:13086–13090.
44. Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature* 395:516–521.
45. Lill, R., and Kispal, G. (2001). Mitochondrial ABC transporters. *Res Microbiol* 152:331–340.
46. Tokatlidis, K., Junne, T., Moes, S., Schatz, G., Glick, B.S., and Kronidou, N. (1996). Translocation arrest of an intramitochondrial sorting signal next to Tim11 at the inner-membrane import site. *Nature* 384:585–588.
47. Arnold, I., Bauer, M.F., Brunner, M., Neupert, W., and Stuart, R.A. (1997). Yeast mitochondrial F1F0-ATPase: the novel subunit e is identical to Tim11. *FEBS Lett* 411:195–200.

48. Leuenberger, D., Bally, N.A., Schatz, G., and Koehler, C.M. (1999). Different import pathways through the mitochondrial intermembrane space for inner membrane proteins. *EMBO J* 17:4816–4822.
49. Claypool, S.M., McCaffery, J.M., and Koehler, C.M. (2006). Mitochondrial mislocalization and altered assembly of a cluster of Barth syndrome mutant tafazzins. *J Cell Biol* 174:379–390.
50. Brandner, K., Mick, D.U., Frazier, A.E., Taylor, R.D., Meisinger, C., and Rehling, P. (2005). Taz1, an outer mitochondrial membrane protein, affects stability and assembly of inner membrane protein complexes: implications for Barth syndrome. *Mol Biol Cell* 16:5202–5214.
51. Paschen, S.A., and Neupert, W. (2001). Protein import into mitochondria. *IUBMB Life* 52:101–112.
52. Koehler, C.M. (2004). The small Tim proteins and the twin Cx3C motif. *Trends Biochem Sci* 29:1–4.
53. Allen, S., Lu, H., Thornton, D., and Tokatlidis, K. (2003). Juxtaposition of the two distal ''CX3C'' motifs via intrachain disulfide bonding is essential for the folding of Tim10. *J Biol Chem* 278:38505–38513.
54. Bauer, M.F., Hofmann, S., and Neupert, W. (2002). Import of mitochondrial proteins. *Int Rev Neurobiol* 53:57–90.
55. Roesch, K., Curran, S.P., Tranebjaerg, L., and Koehler, C.M. (2002). Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a-TIMM13 complex. *Hum Mol Genet* 11:477–486.
56. Hofmann, S., Rothbauer, U., Muhlenbein, N., Neupert, W., Gerbitz, K.D., Brunner, M., and Bauer, M.F. (2002). The C66W mutation in the deafness dystonia peptide 1 (DDP1) affects the formation of functional DDP1. TIM13 complexes in the mitochondrial intermembrane space. *J Biol Chem* 277:23287–23293.
57. Paschen, S.A., Rothbauer, U., Kaldi, K., Bauer, M.F., Neupert, W., and Brunner, M. (2000). The role of the TIM8–13 complex in the import of Tim23 into mitochondria. *EMBO J* 19:6392–6400.
58. Adam, A., Endres, M., Sirrenberg, C., Lottspeich, F., Neupert, W., and Brunner, M. (1999). Tim9, a new component of the TIM22. 54 translocase in mitochondria. *EMBO J* 18:313–319.
59. Koehler, C.M., Leuenberger, D., Merchant, S., Renold, A., Junne, T., and Schatz, G. (1999). Human deafness dystonia syndrome is a mitochondrial disease. *Proc Natl Acad Sci USA* 96:2141–2146.
60. Koehler, C.M., Merchant, S., Oppliger, W., Schmid, K., Jarosch, E., Dolfini, L., Junne, T., Schatz, G., and Tokatlidis, K. (1998). Tim9p, an essential partner subunit of Tim10p for the import of mitochondrial carrier proteins. *EMBO J* 17:6477–6486.
61. Murphy, M.P., Leuenberger, D., Curran, S.P., Oppliger, W., and Koehler, C.M. (2001). The essential function of the small Tim proteins in the TIM22 import pathway does not depend on formation of the soluble 70-kilodalton complex. *Mol Cell Biol* 21:6132–6138.
62. Rothbauer, U., Hofmann, S., Muhlenbein, N., Paschen, S.A., Gerbitz, K.D., Neupert, W., Brunner, M., and Bauer, M.F. (2001). Role of the deafness dystonia peptide 1 (DDP1) in import of human Tim23 into the inner membrane of mitochondria. *J Biol Chem* 276:37327–37334.
63. Sirrenberg, C., Endres, M., Folsch, H., Stuart, R.A., Neupert, W., and Brunner, M. (1998). Carrier protein import into mitochondria mediated by the intermembrane proteins Tim10/Mrs11 and Tim12/Mrs5. *Nature* 391:912–915.
64. Lutz, T., Neupert, W., and Herrmann, J.M. (2003). Import of small Tim proteins into the mitochondrial intermembrane space. *EMBO J* 22:4400–4408.

65. Mackay, J.P., and Crossley, M. (1998). Zinc fingers are sticking together. *Trends Biochem Sci* 23:1–4.
66. Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *EMBO J* 23:3735–3746.
67. Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J.M. (2005). A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. *Cell* 121:1059–1069.
68. Terziyska, N., Lutz, T., Kozany, C., Mokranjac, D., Mesecke, N., Neupert, W., Herrmann, J.M., and Hell, K. (2005). Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions. *FEBS Lett* 579:179–184.
69. Rissler, M., Wiedemann, N., Pfannschmidt, S., Gabriel, K., Guiard, B., Pfanner, N., and Chacinska, A. (2005). The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins. *J Mol Biol* 353:485–492.
70. Allen, S., Balabanidou, V., Sideris, D.P., Lisowsky, T., and Tokatlidis, K. (2005). Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome *c*. *J Mol Biol* 353:937–944.
71. Naoe, M., Ohwa, Y., Ishikawa, D., Ohshima, C., Nishikawa, S.I., Yamamoto, H., and Endo, T. (2004). Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *J Biol Chem* 279:47815–47821.
72. Webb, C.T., Gorman, M.A., Lazarou, M., Ryan, M.T., and Gulbis, J.M. (2006). Crystal structure of the mitochondrial chaperone TIM9. 10 reveals a six-bladed alpha-propeller. *Mol Cell* 21:123–133.
73. Vial, S., Lu, H., Allen, S., Savory, P., Thornton, D., Sheehan, J., and Tokatlidis, K. (2002). Assembly of Tim9 and Tim10 into a functional chaperone. *J Biol Chem* 277:36100–36108.
74. Dekker, P.J., Martin, F., Maarse, A.C., Bomer, U., Muller, H., Guiard, B., Meijer, M., Rassow, J., and Pfanner, N. (1997). The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. *EMBO J* 16:5408–5419.
75. Wiedemann, N., Truscott, K.N., Pfannschmidt, S., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Biogenesis of the protein import channel Tom40 of the mitochondrial outer membrane: Intermembrane space components are involved in an early stage of the assembly pathway. *J Biol Chem* 279:18188–18194.
76. Hoppins, S.C., and Nargang, F.E. (2004). The Tim8-Tim13 complex of *Neurospora crassa* functions in the assembly of proteins into both mitochondrial membranes. *J Biol Chem* 279:12396–12405.
77. Taylor, R.D., and Pfanner, N. (2004). The protein import and assembly machinery of the mitochondrial outer membrane. *Biochim Biophys Acta* 1658:37–43.
78. Gentle, I.E., Burri, L., and Lithgow, T. (2005). Molecular architecture and function of the Omp85 family of proteins. *Mol Microbiol* 58:1216–1225.
79. Tatsuta, T., Model, K., and Langer, T. (2005). Formation of membrane-bound ring complexes by prohibitins in mitochondria. *Mol Biol Cell* 16:248–259.
80. Koehler, C.M., Jarosch, E., Tokatlidis, K., Schmid, K., Schweyen, R.J., and Schatz, G. (1998). Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* 279:369–373.
81. Bauer, M.F., Hofmann, S., Neupert, W., and Brunner, M. (2000). Protein translocation into mitochondria: the role of TIM complexes. *Trends Cell Biol* 10:25–31.
82. Jarosch, E., Tuller, G., Daum, G., Waldherr, M., Voskova, A., and Schweyen, R.J. (1996). Mrs5p, an essential protein of the mitochondrial intermembrane space, affects protein import into yeast mitochondria. *J Biol Chem* 271:17219–17225.

83. Kerscher, O., Holder, J., Srinivasan, M., Leung, R.S., and Jensen, R.E. (1997). The Tim54p-Tim22p complex mediates insertion of proteins into the mitochondrial inner membrane. *J Cell Biol* 139:1663–1675.
84. Koehler, C.M., Murphy, M.P., Bally, N., Leuenberger, D., Oppliger, W., Dolfini, L., Junne, T., Schatz, G., and Or, E. (2000). Tim18p, a novel subunit of the inner membrane complex that mediates protein import into the yeast mitochondrial inner membrane. *Mol Cell Biol* 20:1187–1193.
85. Kerscher, O., Sepuri, N.B., and Jensen, R.E. (2000). Tim18p is a new component of the Tim54p-Tim22p translocon in the mitochondrial inner membrane. *Mol Biol Cell* 11:103–116.
86. Dunn, C.D., and Jensen, R.E. (2003). Suppression of a defect in mitochondrial protein import identifies cytosolic proteins required for viability of yeast cells lacking mitochondrial DNA. *Genetics* 165:35–45.
87. Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H.E., Kuhlbrandt, W., Wagner, R., Truscott, K.N., and Pfanner, N. (2003). Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science* 299:1747–1751.
88. Mohr, J., and Mageroy, K. (1960). Sex-linked deafness of a possibly new type. *Acta Genet Stat Med* 10:54–62.
89. Jin, H., May, M., Tranebjaerg, L., Kendall, E., Fontan, G., Jackson, J., Subramony, S.H., Arena, F., Lubs, H., Smith, S., Stevenson, R., Schwartz, C., *et al.* (1996). A novel X-linked gene, DDP, shows mutations in families with deafness (DFN-1), dystonia, mental deficiency and blindness. *Nat Genet* 14:177–180.
90. Tranebjaerg, L., Schwartz, C., Eriksen, H., Andreasson, S., Ponjavic, V., Dahl, A., Stevenson, R.E., May, M., Arena, F., and Barker, D. (1995). A new X linked recessive deafness syndrome with blindness, dystonia, fractures, and mental deficiency is linked to Xq22. *J Med Genet* 32:257–263.
91. Luft, R. (1994). The development of mitochondrial medicine. *Proc Natl Acad Sci USA* 91:8731–8738.
92. Calvo, S., Jain, M., Xie, X., Sheth, S.A., Chang, B., Goldberger, O.A., Spinazzola, A., Zeviani, M., Carr, S.A., and Mootha, V.K. (2006). Systematic identification of human mitochondrial disease genes through integrative genomics. *Nat Genet* 38:576–582.
93. Steinmetz, L.M., Scharfe, C., Deutschbauer, A.M., Mokranjac, D., Herman, Z.S., Jones, T., Chu, A.M., Giaever, G., Prokisch, H., Oefner, P.J., and Davis, R.W. (2002). Systematic screen for human disease genes in yeast. *Nat Genet* 31:400–404.
94. Aguirre, L.A., del Castillo, I., Macaya, A., Meda, C., Villamar, M., Moreno-Pelayo, M.A., and Moreno, F. (2006). A novel mutation in the gene encoding TIMM8a, a component of the mitochondrial protein translocase complexes, in a Spanish familial case of deafness-dystonia (Mohr-Tranebjaerg) syndrome. *Am J Med Genet A* 140:392–397.
95. Binder, J., Hofmann, S., Kreisel, S., Wohrle, J.C., Bazner, H., Krauss, J.K., Hennerici, M.G., and Bauer, M.F. (2003). Clinical and molecular findings in a patient with a novel mutation in the deafness-dystonia peptide (DDP1) gene. *Brain* 126:1814–1820.
96. Swerdlow, R.H., and Wooten, G.F. (2001). A novel deafness/dystonia peptide gene mutation that causes dystonia in female carriers of Mohr-Tranebjaerg syndrome. *Ann Neurol* 50:537–540.
97. Ujike, H., Tanabe, Y., Takehisa, Y., Hayabara, T., and Kuroda, S. (2001). A family with X-linked dystonia-deafness syndrome with a novel mutation of the DDP gene. *Arch Neurol* 58:1004–1007.
98. Tranebjaerg, L., Hamel, B.C., Gabreels, F.J., Renier, W.O., and Van Ghelue, M. (2000). A *de novo* missense mutation in a critical domain of the X-linked DDP gene causes the typical deafness-dystonia-optic atrophy syndrome. *Eur J Hum Genet* 8:464–467.

99. Roesch, K., Hynds, P.J., Varga, R., Tranebjaerg, L., and Koehler, C.M. (2004). The calcium-binding aspartate/glutamate carriers, citrin and Aralar1, are new substrates for the DDP1/TIMM8a-TIMM13 complex. *Hum Mol Genet* 13:2101–2111.
100. Merchant, S.N., McKenna, M.J., Nadol, J.B., Kristiansen, A.G., Tropitzsch, A., Lindal, S., and Tranebjaerg, L. (2001). Temporal bone histopathologic and genetic studies in Mohr-Tranebjaerg Syndrome (DFN-1). *Otol Neurotol* 22:506–511.
101. Del Arco, A., Agudo, M., and Satrustegui, J. (2000). Characterization of a second member of the subfamily of calcium-binding mitochondrial carriers expressed in human non-excitable tissues. *Biochem J* 345:725–732.
102. Cavero, S., Vozza, A., del Arco, A., Palmieri, L., Villa, A., Blanco, E., Runswick, M.J., Walker, J.E., Cerdan, S., Palmieri, F., and Satrustegui, J. (2003). Identification and metabolic role of the mitochondrial aspartate-glutamate transporter in *Saccharomyces cerevisiae*. *Mol Microbiol* 50:1257–1269.
103. Lasorsa, F.M., Pinton, P., Palmieri, L., Fiermonte, G., Rizzuto, R., and Palmieri, F. (2003). Recombinant expression of the Ca(2+)-sensitive aspartate/glutamate carrier increases mitochondrial ATP production in agonist-stimulated Chinese hamster ovary cells. *J Biol Chem* 278:38686–38692.
104. Ramos, M., del Arco, A., Pardo, B., Martinez-Serrano, A., Martinez-Morales, J.R., Kobayashi, K., Yasuda, T., Bogonez, E., Bovolenta, P., Saheki, T., and Satrustegui, J. (2003). Developmental changes in the $Ca^{2+}$-regulated mitochondrial aspartate-glutamate carrier aralar1 in brain and prominent expression in the spinal cord. *Brain Res Dev Brain Res* 143:33–46.
105. Begum, L., Jalil, M.A., Kobayashi, K., Iijima, M., Li, M.X., Yasuda, T., Horiuchi, M., del Arco, A., Satrustegui, J., and Saheki, T. (2002). Expression of three mitochondrial solute carriers, citrin, Aralar1 and ornithine transporter, in relation to urea cycle in mice. *Biochim Biophys Acta* 1574:283–292.
106. Palmieri, L., Pardo, B., Lasorsa, F.M., del Arco, A., Kobayashi, K., Iijima, M., Runswick, M.J., Walker, J.E., Saheki, T., Satrustegui, J., and Palmieri, F. (2001). Citrin and Aralar1 are $Ca^{2+}$-stimulated aspartate/glutamate transporters in mitochondria. *EMBO J* 20:5060–5069.
107. Kobayashi, K., Sinasac, D.S., Iijima, M., Boright, A.P., Begum, L., Lee, J.R., Yasuda, T., Ikeda, S., Hirano, R., Terazono, H., Crackower, M.A., Kondo, I., *et al.* (1999). The gene mutated in adult-onset type II citrullinaemia encodes a putative mitochondrial carrier protein. *Nat Genet* 22:159–163.
108. James, A.M., Sheard, P.W., Wei, Y.H., and Murphy, M.P. (1999). Decreased ATP synthesis is phenotypically expressed during increased energy demand in fibroblasts containing mitochondrial tRNA mutations. *Eur J Biochem* 259:462–469.
109. Davey, K.M., Parboosingh, J.S., McLeod, D.R., Chan, A., Casey, R., Ferreira, P., Snyder, F.F., Bridge, P.J., and Bernier, F.P. (2006). Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCM A syndrome, a novel autosomal recessive Barth syndrome-like condition. *J Med Genet* 43:385–393.
110. Silvestri, L., Caputo, V., Bellacchio, E., Atorino, L., Dallapiccola, B., Valente, E.M., and Casari, G. (2005). Mitochondrial import and enzymatic activity of PINK1 mutants associated to recessive parkinsonism. *Hum Mol Genet* 14:3477–3492.
111. Devi, L., Prabhu, B.M., Galati, D.F., Avadhani, N.G., and Anandatheerthavarada, H.K. (2006). Accumulation of amyloid precursor protein in the mitochondrial import channels of human Alzheimer's disease brain is associated with mitochondrial dysfunction. *J Neurosci* 26:9057–9068.

This page intentionally left blank

# 15

# *The Role of the TIM23 Complex and Its Associated Motor Complex in Mitochondrial Protein Import*

JAN DUDEK[a] • BERNARD GUIARD[b] • PETER REHLING[a]

[a]*Institut für Biochemie und Molekularbiologie*
*Universität Freiburg*
*Hermann-Herder-Straße 7*
*D-79104 Freiburg, Germany*

[b]*Centre de Génétique Moléculaire du CNRS*
*F-91190 Gif-sur-Yvette, France*

## I. Abstract

Mitochondrial matrix proteins utilize N-terminal presequences to direct them to their destination within the organelle. After crossing the outer mitochondrial membrane, the presequence targets precursor proteins to a specific translocase in the inner membrane for further translocation: the presequence translocase or TIM23 complex. The presequence translocase forms a pore in the inner membrane through which precursors are transported toward the matrix. A minor fraction of presequence proteins are not completely transported into the matrix, but become laterally integrated into the inner membrane. Two distinct states of the translocase seem to function in these two different transport pathways: the $TIM23^{SORT}$ complex integrates membrane proteins in a membrane potential-dependent manner. For the translocation of a matrix protein, the $TIM23^{PAM}$ complex binds the presequence translocase-associated motor complex (PAM), which is



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25015-2

necessary to provide additional driving force for the full translocation across the inner membrane.

## II. Introduction

Two lipid bilayers divide mitochondria into four subcompartments. The outer and inner mitochondrial membranes provide a partition for two aqueous compartments, the mitochondrial matrix and mitochondrial intermembrane space (IMS). Each of the four compartments is resident to an elaborate collection of proteins that fulfill the various functions of the mitochondrion. Since the mitochondrial genome encodes only a small number of proteins, the majority of mitochondrial proteins are encoded by the nuclear genome, translated on cytosolic ribosomes, and kept in an unfolded and thus transport competent state by cytosolic chaperones [1–7]. An elaborate system of translocases and assembly machineries within the organelle are necessary to direct nuclear-encoded precursors to their final destination.

Protein transport across the outer membrane is mediated by the translocase of the outer membrane (TOM) complex, which serves as the general entry gate into the organelle (Figure 15.1) (see Chapter 14 and [1, 3, 4, 8]). Following translocation through the TOM complex, protein transport pathways diverge depending on the targeting signals in the precursor proteins, which dictate their final intramitochondrial location. Outer membrane proteins of the β-barrel family are initially translocated across the outer membrane through the TOM complex. Subsequently, they contact the small Tim proteins in the IMS, which deliver β-barrel precursors to the sorting and assembly machinery (SAM) complex for integration into the outer membrane (Chapter 14). Proteins of the IMS can be released from the TOM translocase and either fold spontaneously into their native conformation or are assisted in folding and assembly by specialized factors in the IMS such as the Mia40-Erv1 system. Proteins of the carrier family are transported from the TOM complex with the aid of the small Tim proteins to the inner membrane carrier translocase (TIM22 complex) for inner membrane insertion (Chapter 16). Finally, a large number of mitochondrial proteins carry N-terminal targeting signals, so-called presequences, which direct them across the outer membrane via the TOM complex and eventually into or across the inner mitochondrial membrane [1–4, 8]. On translocation through the TOM complex, presequence bearing precursors are directed to the presequence translocase or TIM23 complex (Figure 15.1). The TIM23 complex generally directs precursor proteins into the mitochondrial matrix. However, in the presence of additional targeting/sorting

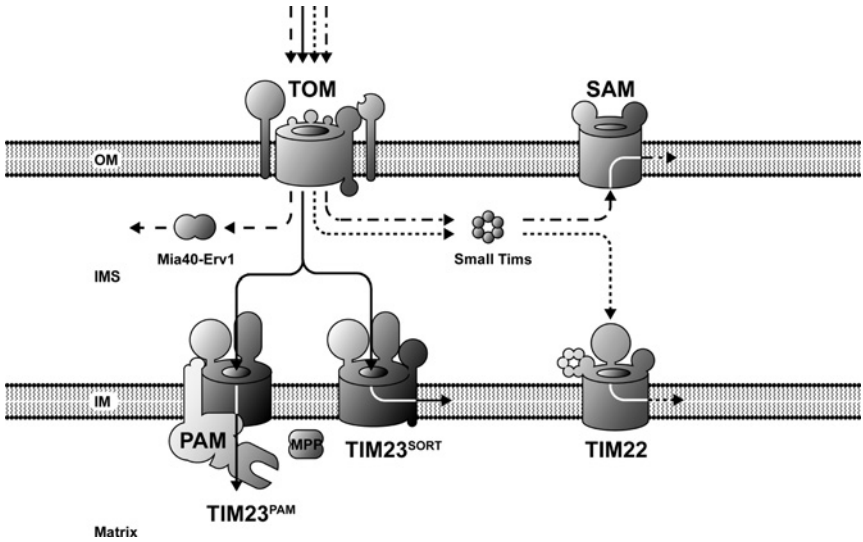



FIG. 15.1. Protein-sorting pathways in the mitochondrion. Proteins cross the outer mitochondrial membrane (OM) through the TOM complex. Proteins with internal targeting sequences are released into the IMS, become integrated into the outer membrane by the SAM, or become integrated into the inner membrane (IM) by the TIM22 complex. Proteins that are targeted into mitochondria by a presequence also cross the outer membrane through the TOM complex and are then transported by the presequence translocase (TIM23), which exists in two forms. Most presequence proteins are transported into the matrix by the $TIM23^{PAM}$ complex, which acts together with the PAM complex. A different composition of this translocase ($TIM23^{SORT}$) is needed to integrate presequence proteins into the inner membrane.

information in the transported polypeptide chain, proteins can also be inserted into the inner membrane where they may remain as integral membrane proteins or may be released into the IMS after proteolytic processing steps [1–3, 5]. The presequence translocase forms a tightly regulated aqueous pore in the inner membrane through which the preproteins are translocated. The mitochondrial membrane potential ($\Delta\psi$) drives the initial translocation of the positively charged presequence across the inner membrane and provides a sufficient driving force to promote lateral insertion of precursors into the inner membrane by the TIM23 complex [4, 8, 9]. Yet, the majority of presequence proteins are soluble proteins of the mitochondrial matrix. These proteins use the presequence translocase for complete translocation across the inner membrane. While the $\Delta\psi$ serves to drive transport of the presequence of such precursors across the inner membrane, it is insufficient for full precursor translocation. Further transport of the precursor requires cooperation of the presequence translocase-associated motor complex (PAM).

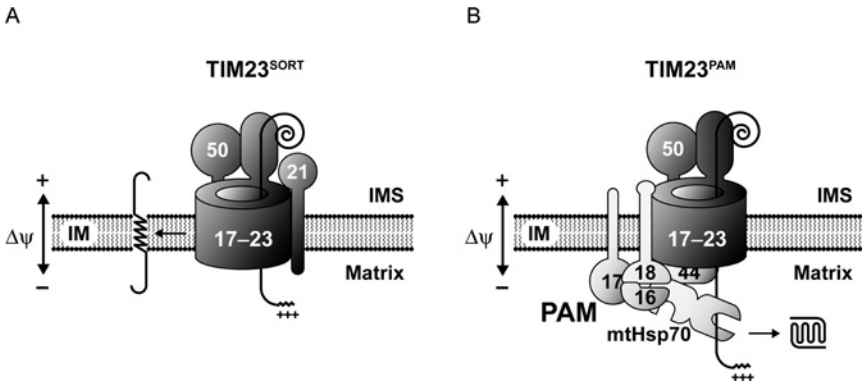



FIG. 15.2. Presequence proteins with different destinations in the mitochondrion are transported by two different forms of the presequence translocase. Precursor proteins with sorting signals are transported in a membrane potential ($\Delta\psi$)-dependent manner by the $TIM23^{SORT}$ complex. The $TIM23^{SORT}$ complex consists of Tim23, Tim17, Tim50, and Tim21. For the membrane potential-dependent transport of proteins into the matrix, Tim21 is removed and the $TIM23^{PAM}$ complex associates with the PAM complex to drive the complete translocation of preproteins across the inner membrane. The PAM complex consists of the motor protein mtHsp70 and the components Tim44, Pam16, Pam17, and Pam18.

This specialized ATP-driven complex is recruited to the pore-forming TIM23 complex to allow matrix transport (Figure 15.2) [2, 3, 10–13]. On import into the matrix, the precursor proteins are processed by the mitochondrial processing peptidase (MPP) into their mature form [14]. Molecular chaperones in the matrix help to fold the mature proteins into their native conformation [1, 15].

## III. Mitochondrial Presequence Proteins

Proteins are usually targeted to different organelles by intrinsic signals. In a number of cases, such as for many mitochondrial proteins, the targeting signal is located at the N-terminus of the preprotein. These signal sequences are recognized by specific receptors on the surface of the organelle permitting subsequent translocation of preproteins into the organelle. Usually the signal sequences do not contribute to the folding and function of a protein and are cleaved by specific signal peptidases [16, 17]. In the baker's yeast *Saccharomyces cerevisiae*, it has been estimated that at least 43% of nuclear-encoded mitochondrial precursors are targeted to the organelle by an N-terminal presequence [18].

## A. The Structure of the Presequence

Mitochondrial presequences are typically amphipathic α-helical segments formed by a stretch of 20–80 amino acids, which consists of hydrophobic, hydroxylated, or basic amino acids and have a clear bias against acidic residues [1, 19–21]. The targeting information is not encoded by a consensus sequence but rather by the distinct secondary structure and charge distribution. Thus, natural presequences can be replaced with more simplified amino acid sequences as long as this sequence resembles the secondary structure and charge pattern of natural presequences [1, 8]. Mitochondrial presequences have been shown to be necessary and sufficient for the transport of proteins into the organelle. A presequence fused to the N-terminus of a nonmitochondrial protein can successfully target the foreign protein into mitochondria, indicating that it contains all necessary information for mitochondrial import [1, 8]. Accordingly, proteins are no longer imported into mitochondria when the presequence has been removed.

## B. Functions of the Presequence

The distinct structural features of the presequence play important roles at several steps during import. Initially, recognition of the presequence protein occurs by interaction of the presequence with mitochondrial import receptors. The primary import receptor Tom20, a component of the TOM complex, forms a hydrophobic binding pocket that accommodates the hydrophobic side of the α-helical presequence [22, 23]. After initial recognition of the presequence by Tom20, other receptors of the TOM complex execute their roles in the recognition of preproteins. These include the receptor domains of Tom22, the major receptor of the TOM complex, and Tom5, which has a function in the transfer of preproteins to the channel-forming unit of the TOM complex, Tom40. Thus, these components form a chain of binding sites for presequences during transport across the outer membrane [1, 3, 8, 24], promoting the movement of preproteins across the outer membrane and promoting specificity in the import pathway. These interactions are thought to be mostly ionic and thus based on the charged amino acids of the presequence. For translocation across the inner membrane, these charges are similarly important. At the inner membrane, the receptor domain of the inner membrane protein Tim23, the pore-forming subunit of the TIM23 complex, recognizes the presequence [25–27]. Moreover, the $\Delta\psi$ generated by the activity of the respiratory chain complexes exerts an electrophoretic force on the positively charged residues of the presequence. As it will be discussed later, this force contributes

to the inward directed motion of preproteins across the inner membrane [24, 28, 29].

### C. Processing and Sorting of the Mature Protein

Although presequences play a vital role in early targeting processes, they are usually not part of the mature protein since the presequence can interfere with the folding of proteins into their native conformations [30]. On entry of the presequence into the mitochondrial matrix, it is usually removed proteolytically by the MPP [14, 31]. Some proteins undergo a second processing in the matrix or IMS after MPP cleavage. Such proteins are processed by the mitochondrial intermediate peptidase (MIP) or inner membrane peptidase (IMP) [14, 32]. It is only in a few rare cases that the signal sequence remains part of the mature protein [33–35].

While the presequence directs proteins into the matrix, additional targeting information within the N-terminus leads to transport into the inner membrane. A hydrophobic sorting sequence located behind the presequence stalls transport across the inner membrane and leads to lateral integration of the protein into the membrane, a process referred to as inner membrane sorting [3–5, 8]. In some proteins, the second signal sequence resembles bacterial signal sequences and consists of a hydrophobic stretch preceded by basic amino acids [1, 4, 8]. Proteins with such a bipartite presequence are also integrated into the inner membrane or can be exposed to a further processing event mediated by the IMP, a protein complex with its catalytic domain facing the IMS, resulting in the generation of a soluble IMS protein [32, 36].

## IV. The Presequence Translocase: TIM23 Complex

The presequence translocase or TIM23 complex selectively translocates presequence containing preproteins across the inner membrane. Integrated into the inner membrane, it consists of a central pore-forming core, through which preproteins are translocated across the membrane. The essential Tim23 protein represents the actual pore-forming subunit and contains a receptor domain in the IMS that recognizes the presequence of preproteins [26, 27]. While the TIM23 complex alone is sufficient for sorting of proteins into the inner membrane, translocation of precursors into the matrix additionally requires recruitment of the PAM complex.

## A. Components of the TIM23 Complex

The TIM23 complex consists of three essential membrane proteins: Tim23, Tim17, and Tim50 (Figure 15.2). The first component of the presequence translocase that was identified by a genetic screen for yeast mutants with a transport defect for mitochondrial proteins was Tim23 [37, 38]. Tim23 is predicted to form four α-helical transmembrane domains with its C-terminal half, while an N-terminal domain of ~100 amino acids is exposed into the IMS. The same genetic screen led to discovery of Tim17 as a further component of the presequence translocase [39]. In addition, Tim17 has been found as a multicopy suppressor of a temperature-sensitive *tim23* mutant [40]. The sequence of Tim17 shows similarity to the C-terminal membrane integral part of Tim23, but lacks the extended N-terminal domain. Interestingly, the four transmembrane helices of Tim23 and Tim17 display sequence similarities to a third mitochondrial protein, involved in protein translocation, Tim22. Tim22 is the core component of the TIM22 complex, which integrates carrier proteins into the inner membrane. It has been suggested that Tim17, Tim23, and Tim22 represent a novel family of transporter proteins. Interestingly, two amino acid transporters, one of the bacterial plasma membrane and one of the outer chloroplast envelope, belong to this protein family. Thus, it has been speculated that Tim22, Tim23, and Tim17 have been derived from these transporters [41]. Tim17 and Tim23 are present in equal amounts in the TIM23 complex and a stable interaction between both protein has been shown by different techniques [42–45]. Isolation of the Tim23-Tim17 complex and a cross-linking approach led to the identification of the third essential component of the presequence translocase complex, Tim50 [46–48]. Tim50 is integrated into the membrane by a single transmembrane helix and exposes a C-terminal domain to the IMS (Figure 15.2). The IMS domain of Tim50 ($Tim50_{IMS}$) interacts with the N-terminal domain of Tim23 [46, 47]. Similarly, to the other essential components of the TIM23 complex, Tim50 is involved in the translocation of preproteins and can be chemically cross-linked to precursor proteins arrested during transport, indicating that Tim23, Tim17, and Tim50 are in proximity to the translocating polypeptide chain [46–50]. While cross-links of precursors to Tim23 and Tim17 have been observed during translocation in energized mitochondria, a cross-link of a precursor to Tim50 is also found in mitochondria in which the membrane potential has been dissipated [47, 48]. Moreover, Tim50 promotes association of the precursors protein with the TOM complex [13]. Thus, Tim50 acts in an early phase of the transport process.

Affinity purification of the intact TIM23 complex led to identification of Tim21, a further component of the presequence translocase [13, 51]. Tim21

is integrated in the inner membrane with its N-terminus. The C-terminal hydrophilic domain is exposed to the IMS. The function of Tim21 will be discussed later in greater detail.

### B. The Tim23 Channel and Its Regulation

In electrophysiological analyses, Tim23 was identified as the pore-forming component of the presequence translocase [27]. A channel with similarity to Tim23 had already been detected in previous electrophysiological studies of inner mitochondrial membranes; however, its molecular identity had remained open [52]. The Tim23 channel is selective for cations and is activated by presequences [27]. While the C-terminus forms the actual channel, the IMS domain is critical for presequence recognition and thus regulation of the channel [26, 27]. The sequence similarity between Tim17 and Tim23 has led to the proposal that Tim17 may also form a channel in the inner membrane; however, experimental evidence in support of this idea is still lacking. Based on electrophysiological analysis and import experiments using particles, the size of the Tim23 pore has been estimated to be about 1.3 nm [27, 53]. This corresponds to the pore size of the carrier translocase, which has a diameter of about 1.2–1.7 nm [54, 55]. However, it is smaller than the pore of the TOM complex (about 2.2 nm) [56, 57].

While necessary for protein transport, protein-translocating pores in the inner mitochondrial membrane pose a certain problem for mitochondrial function. Mitochondrial respiratory chain complexes convert metabolic energy into a proton gradient across the inner membrane. This gradient drives the $F_oF_1$-ATPase to generate the majority of cellular ATP. Therefore, the inner mitochondrial membrane has to be tightly sealed to avoid leakage of protons and thus collapse of the membrane potential. Thus, a pore that is capable of transporting preproteins must be tightly regulated such that channel opening only takes place for the transport of preproteins. It became clear that the IMS domains of Tim50 and Tim23 cooperate in this process [26, 58]. Closure of the Tim23 channel is mediated through association of $Tim50_{IMS}$ with $Tim23_{IMS}$ [58]. Opening of the channel is initiated by the presequence of the incoming preprotein, probably via contact with the IMS domain of Tim23 [26, 27, 58]. The antagonistic action of Tim50 and preproteins on the channel allow for on demand opening of the channel while maintaining the inner membrane permeability barrier in the inactive state [58]. Additionally, a function for Tim17 in the regulation of the pore has been proposed based on the analysis of Tim17 mutant phenotypes [59].

## V. Energy Requirement for Matrix Translocation: The Motor Complex

### A. The Membrane Potential and ATP-Hydrolysis Drive Protein Import

Transport of proteins across the inner mitochondrial membrane depends on the $\Delta\psi$ irrespective of which of the two TIM translocases is used [1–5, 8, 60]. As described above, the $\Delta\psi$ exerts an electrophoretic force on the charged presequences and thus helps to move the N-terminal part of the precursor proteins across the membrane toward the matrix. Moreover, the $\Delta\psi$ is important for activation of the Tim23 channel [26, 27]. However, due to the spatial restriction of the $\Delta\psi$, it does not provide sufficient driving force for full translocation of proteins into the matrix. Thus, a second driving force is necessary for completion of the translocation process. In addition to the $\Delta\psi$, transport of matrix proteins requires hydrolysis of ATP by the PAM complex [2, 10, 12, 61–64]. The actual ATPase of this complex is a chaperone of the family of Hsp70 proteins, mtHsp70 [65]. In addition to mtHsp70, the PAM complex consists of five regulatory components Tim44, Mge1, Pam16, Pam17, and Pam18 that are described in the following sections (Figure 15.2).

### B. MtHsp70: The Central Component of the Motor Complex

Chaperones of the Hsp70 family exert a broad spectrum of functions within the cell, including assistance in folding of newly synthesized or misfolded proteins and prevention or reversion of their aggregation [66–68]. Hsp70 proteins consist of an ATPase domain and a substrate-binding domain. In an ATP-dependent manner, the substrate-binding domain switches between a closed conformation that tightly binds unfolded substrate proteins and an open state, in which bound substrates can be released. Hydrolysis of ATP leads to closure of the substrate-binding domain, whereas the exchange of ADP against ATP leads to its reopening [66, 67, 69]. The low intrinsic ATPase activity of Hsp70 is stimulated by cochaperones of the J-protein family [70–73]. Exchange of nucleotides is assisted by cochaperones, which belong to the class of the GrpE protein family [66, 69].

Important insights into the function of mtHsp70 in protein translocation were obtained through the analysis of mitochondria with mutant forms of the protein, which displayed strong import defects. While in certain mutants this defect can be partially overcome by chemical unfolding of the precursor prior to import, other mtHsp70 mutants display a full import defect for all forms of

matrix preproteins [65, 74]. In addition, mtHsp70 mutant mitochondria display defects in folding of newly imported proteins in the mitochondrial matrix [65]. This phenotype suggests that mtHsp70 is necessary for the transport of proteins. It unfolds these proteins during import, promotes translocation across the membranes, and subsequently assists in folding into their native conformation [10, 12, 15]. For folding of proteins in the matrix, mtHsp70 forms a soluble complex with its substrate proteins and its activity is controlled by the cochaperones Mge1 and Mdj1 [15]. For the transport of proteins across the membrane, Hsp70 is part of the import motor complex PAM, which is bound to the presequence translocase [2, 9, 12, 13, 75].

### C. Components of the Motor Complex

Besides mtHsp70, additional components of the PAM complex have been identified (Figure 15.2). The genetic screen that identified Tim23 and Tim17 as components of the mitochondrial translocase also led to the identification of a further essential component Tim44 [76]. Tim44 can be chemically cross-linked to precursor proteins during translocation and plays an important role in protein transport [77–80]. Tim44 does not contain a classical transmembrane domain, but is peripherally attached to the inner membrane from the matrix side [43]. While Tim44 is able to directly interact with the membrane, it has also been shown to interact with Tim23 [81, 82]. MtHsp70 associates with Tim44 in an ATP-dependent manner and binding of ATP to mtHsp70 leads to dissociation of Tim44 from mtHsp70 *in organello* [83–85]. The current view is that Tim44 acts as an anchor for the binding of mtHsp70 to the presequence translocase. Association of mtHsp70 to Tim44 is thought to be mainly mediated via the ATPase domain; however, additional binding sites in the peptide-binding domain seem also to contribute to the interaction between the two proteins [12, 86–88]. Besides the Tim44-mediated recruitment of mtHsp70 to the presequence translocase, mtHsp70 appears to associate to the presequence translocase also via Tim17; however, this interaction is independent of ATP [89].

As the activity of Hsp70 proteins is generally regulated by cochaperones, it has been assumed that in analogy also at the inner membrane cochaperones would regulate mtHsp70 during import of preproteins. The essential exchange factor Mge1 was identified by its strong sequence homology to bacterial GrpE. Mge1 has a function in protein folding as well as in import of proteins into mitochondria [15, 90]. Its role in protein transport was shown by accumulation of mitochondrial precursor proteins after down-regulation of *MGE1*-expression in yeast cells [91–93]. The second cochaperone of mtHsp70 in the motor complex, Pam18 (Tim14), was only

recently identified. Pam18 is an integral inner membrane protein essential for yeast viability [94–96]. Pam18 possesses a single transmembrane segment located in the N-terminal half of the protein, which divides the protein into two domains, one small domain at the extreme N-terminus facing the IMS and a C-terminal domain located in the matrix (Figure 15.2). The N-terminal IMS domain of Pam18 binds to Tim17 and is important for the recruitment of the PAM complex to the presequence translocase [13]. The matrix domain shows strong sequence homology to J-domains and indeed the ATPase activity of mtHsp70 in the presence of the purified Pam18 matrix domain is strongly increased [94, 96]. This indicates that Pam18 is a specific activator of mtHsp70 at the presequence translocase. In agreement with this, a defect in the transport of matrix proteins is observed when the function of Pam18 is compromised [94–96].

Another essential cochaperone of the import motor is Pam16 (Tim16) [34, 35]. Pam16 is peripherally bound to the inner membrane and similarly to Pam18 is required for import into the mitochondrial matrix. Yet, both Pam16 and Pam18 are dispensable for inner membrane sorting, a process that is largely independent of motor function [97]. Pam16 interacts with the J-domain of Pam18 forming an 80-kDa complex [34]. The sequence of Pam16 shows significant homology to J-domains but the critical amino acids ''HPD'' which are crucial for activation of Hsp70 by J-proteins are substituted. Indeed, measurements of the mtHsp70 ATPase activity in the presence of the purified Pam16 did not show any stimulation. Instead, the activation of mtHsp70 by Pam18 was efficiently inhibited by addition of Pam16 [98, 99]. Thus, Pam16 was classified as a novel inhibitory cochaperone that has homologues in various organisms ranging from *Caenorhabditis elegans*, *Drosophila melanogaster*, and humans. Recently, Pam16 was reported to also bind to the nonessential J-protein Mdj2. Interestingly, *in vitro* data suggests that binding of Pam16 to Mdj2 enhances its capability to activate mtHsp70 instead of inhibiting it [100]. But since Mdj2 is not a component of the active translocase, the relevance of this interaction for the transport of mitochondrial proteins remains open.

Another component of the presequence translocase is Pam17, which was identified by affinity purification of the presequence translocase [101]. Pam17 is integrated into the inner membrane by two transmembrane domains and exposes a hydrophilic C-terminal domain to the matrix (Figure 15.2). Deletion of *PAM17* in yeast destabilizes the motor complex and strongly impairs the import of matrix targeted proteins, whereas the sorting of proteins into the inner membrane is not affected [101]. These findings indicate that Pam17 has a function in the structural organization of the motor complex.

## VI. Models of Motor Function

### A. Protein Unfolding for Import

Proteins that are translated in the cytosol will not remain unfolded, but may acquire a partially folded conformation before they reach the outer mitochondrial membrane. This becomes apparent with the use of *in vitro* import assays in which preproteins that have been synthesized in a cell-free translation system are incubated with isolated mitochondria. The import of mitochondrial precursor proteins fused to DHFR is blocked when the ligand methotrexate is added [10, 102]. As binding of a ligand is indicative of protein folding, this finding supports the notion that a protein must be in an unfolded conformation for transport. Moreover, estimations of the pore diameter of the TOM complex and the TIM23 translocase show that the pores are not wide enough for the translocation of precursors in a folded state. Thus, proteins must be unfolded before import into mitochondria [1, 10, 11]. The role of cytosolic chaperones in unfolding of mitochondrial preproteins is still unclear. In *in vitro* import assays, ATP depletion on the outside of mitochondria prevents the import of aggregation prone hydrophobic membrane proteins but not of hydrophilic matrix proteins [103]. Based on these findings, it was suggested that the major role of chaperones in the cytosol is to prevent aggregation rather than to unfold proteins before their import. This implies that mitochondria are able to actively unfold preproteins. Indeed, an active contribution of mitochondria to the unfolding of proteins was observed. Unfolding of proteins in the presence of mitochondria is much faster than unfolding of the same protein in solution [104]. Additionally, mitochondria influence the folding pathway of proteins. The spontaneous unfolding of a protein in solution is a global process in which large parts of the protein unfold at an early stage. In contrast to this, proteins which are unfolded by mitochondria start their unfolding at the N-terminus [105, 106]. The unfolding of a protein is an energy-dependent process. The two energy sources that play a role in protein transport, the $\Delta\psi$ and mtHsp70, both contribute to unfolding of preproteins. In the early stages of import in which the $\Delta\psi$ exerts an electrophoretic force on the presequence, the pulling force also contributes to unravel the N-terminal parts of the protein [105]. In addition, it is clear that the import motor also contributes to the unfolding process, since mtHsp70 mutant mitochondria display defects in the import of folded but not of chemically unfolded proteins [10–12].

### B. Two Models of Hsp70 Function for Matrix Import

Two different models have been proposed to describe how the PAM complex supports the transport of proteins across the inner membrane: the trapping model and the pulling model [12, 87]. In the trapping model or

Brownian ratchet model, the import is driven by the Brownian movement of the precursor itself, which leads to a random in- and outward directed movement in the pore of the presequence translocase. As soon as the unfolded preprotein is moving toward the matrix, it will expose a binding site for mtHsp70. The interaction of mtHsp70 with the polypeptide chain induces ATP-hydrolysis and the tight binding of the chaperone to the substrate protein. This tight binding prevents the backsliding of the imported parts of the precursor protein and the precursor will remain at its position until the next inward directed movement [12, 87]. In the trapping model, the import motor provides the energy to transform the random Brownian movement into a vectorial inward directed motion. Tim44 acts as a molecular anchor that enriches the concentration of mtHsp70 at the import pore and probably positions mtHsp70 at the pore for an efficient interaction with the precursor protein (Figure 15.2). In the pulling or power-stroke model, mtHsp70 drives the import process by exerting a pulling force on the polypeptide chain. MtHsp70 is anchored to Tim44 at the import pore in its open conformation. The binding of a substrate protein and the subsequent ATP-hydrolysis induces the closing of the substrate domain. The pulling model proposes a second conformational change, which alters the position of the substrate-binding domain toward the ATPase domain. If the ATPase domain is stably anchored to Tim44, this conformational change can generate a pulling force on the precursor protein. This force pulls the protein across the membrane into the matrix [12, 15, 87]. In this model, Tim44 not only positions mtHsp70 at the import pore but also has an important function to act as a molecular fulcrum, necessary for the generation of the pulling force. Like the motor proteins myosin and kinesin, mtHsp70 transforms energy derived from the hydrolysis of ATP into a directed movement of proteins [10, 12]. However, the recently emerging picture is that probably both mechanisms may apply during protein import. While loosely folded or unfolded precursors may primarily be imported by the trapping mechanism, it is becoming clear that folded precursors depend on the pulling mechanism [107, 108].

### C. Functional Implication of PAM Organization for Hsp70 Regulation

The PAM complex is emerging as quite a complicated molecular machine that regulates mtHsp70 during the import process. Several analyses and the identification of novel PAM components have shed new light on the steps of the mtHsp70 cycle. *In organello* mtHsp70 binds to Tim44 in an ATP-dependent manner [86, 88, 109]. In contrast, the *in vitro* interaction of the purified proteins is independent of the nucleotide-binding state

of mtHsp70 [110]. Further *in vitro* analyses revealed that although ADP-mtHsp70 binds with the similar affinity to Tim44 as ATP-mtHsp70, its interaction is sensitive to the addition of Mge1, whereas ATP-mtHsp70 is not. This *in vitro* data suggests an unexpected function of Mge1 at the import pore. By releasing nonproductive ADP-mtHsp70 from the import pore, it increases the concentration of the active ATP-mtHsp70 [110, 111].

The PAM complex integrates mtHsp70 in a way that supports efficient interaction of ATP-mtHsp70 with substrate proteins. At the same time, the J-domain of Pam18 activates the ATPase activity in order to allow closure of the substrate-binding domain on the incoming precursor. How can the machinery prevent premature activation of mtHsp70? It appears that a major function of Pam16 is to prevent Pam18 from premature or unproductive activation of mtHsp70. This inhibition is probably overcome only after a substrate protein appears in the import pore and the incoming polypeptide is bound by mtHsp70. The interaction with the substrate activates mtHsp70, but also independent of this changes its affinity for Tim44. As described earlier, *in vitro* experiments with purified proteins show that ATP-mtHsp70 and ADP-mtHsp70 bind to Tim44. However, ATP-mtHsp70-Tim44 complexes are unstable in the presence of preproteins [110]. Thus, it was concluded that during the import of preproteins the interaction of mtHsp70 with substrate proteins induces the release of mtHsp70 from Tim44 and from the motor complex. At this stage, Pam18 activates the ATPase activity and leads to closure of the substrate-binding domain. The strong acceleration of the ATPase activity by Pam18 ensures that mtHsp70 stably interacts with its substrate protein before dissociation from the import pore [112]. Subsequently, the stable interaction of mtHsp70 with the substrate protein is released by action of Mge1. It exchanges the bound nucleotides in mtHsp70 and allows opening of the substrate-binding domain [110].

## VII. Transport of Proteins Across Two Membranes

### A. Structural Connection Between Translocases on the Outer and Inner Membrane

In mitochondria, the distance between the outer and the inner membrane varies. At distinct sites, both membranes come into proximity to each other. These so-called ''contact sites'' have been shown to be the sites of protein translocation across two membranes. A precursor protein has been arrested during its import by binding a specific antibody at its C-terminus. Labeling the arrested precursor with protein A-gold particles and analysis by electron microscopy revealed that most preproteins accumulate at contact sites [102, 113]. The distance between the two closely apposed

membranes has been estimated by electron microscopy to be about 15–18 nm. [102, 114]. In a systematic analysis, the minimal amount of amino acids necessary to span both membranes has been determined to be about 50 in an extended conformation [114]. The presence of contact sites between the outer and inner membrane raised the question if the TOM and TIM translocases directly interact with each other. Interestingly, based on protease protection assays, Tim23 was suggested to span the outer mitochondrial membrane with its N-terminal amino acids [115]. However, an interaction between this domain and the TOM complex was not found and later studies showed that the N-terminus of Tim23 is dispensable for formation of a TOM–TIM supercomplex connected by a translocating precursor protein [116].

Besides Tim50 and Tim23, Tim21 also exposes a domain to the IMS ($Tim21_{IMS}$). Binding studies with the purified $Tim21_{IMS}$ showed that this domain directly interacts with the IMS domain of Tom22 ($Tom22_{IMS}$), the central receptor unit of the TOM complex (Figure 15.3) [13, 51]. The IMS domain of Tom22 is important for protein transport since it binds the presequence of precursor proteins after their passage through the TOM complex [13, 25, 117–120]. Interestingly, the interaction of the presequence with $Tom22_{IMS}$ domain is released on binding of $Tim21_{IMS}$ to $Tom22_{IMS}$, indicating that the presequences and Tim21 use the same binding site [13]. Thus, the emerging picture of the early steps in protein transport from the TOM complex to the TIM23 complex is initiated by association of $Tim21_{IMS}$ to $Tom22_{IMS}$. This interaction releases the presequence from the TOM complex and allows the precursor to contact the core of the presequence translocase for inner membrane translocation (Figure 15.3).

## B. The TOM–TIM23 Supercomplex

Despite the structural connection between the TIM and TOM, several studies show that the presequence translocase can function independent of the TOM complex. After removal of the outer membrane, the presequence translocase is still able to transport precursor proteins [121, 122]. However, for an efficient transport of proteins across both membranes *in vivo*, both translocases need to work in a tightly coupled manner. Indeed, preproteins that are transported through the TOM complex are directly handed over to the presequence translocase and thus accumulation of a free transport intermediate in the IMS has not been detected [1, 8]. During transport across both membranes, the precursor protein spans both TOM complex and the presequence translocase simultaneously in an extended conformation. Experimental evidence for this was obtained through the use of translocation intermediates of precursor proteins with a tight-folded domain at their

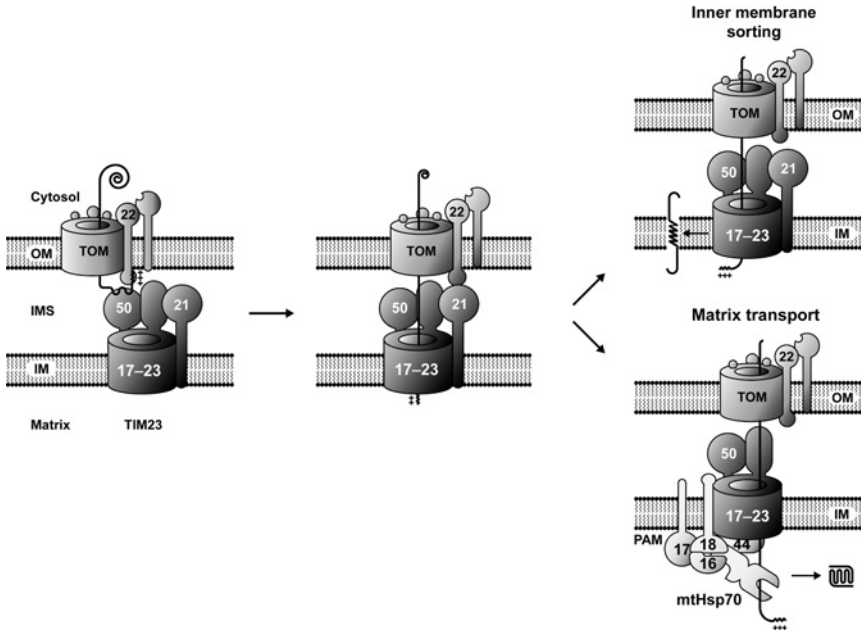



FIG. 15.3 The translocation of a presequence proteins depicted in discrete steps. Presequences that are transported across the outer membrane through the TOM complex bind to Tom22, a core component of the translocase, as soon as they reach the IMS. The presequence protein is chased from this interaction, when Tim21, a component of the presequence translocase, interacts with the same binding site of Tom22. The further translocation of a membrane protein occurs by the $TIM23^{SORT}$ complex, which permits integration of the protein into the inner membrane. For the transport of a matrix protein, the translocase switches to the $TIM23^{PAM}$ form and recruits the import motor PAM, which provides the energy necessary for the complete translocation across the membrane.

C-terminus. The tightly folded domain prevents the complete translocation of the protein through the TOM pore and the arrested protein connects the TOM and the TIM23 complex. This connection is stable enough for affinity purification. The purification of components of the TOM complex also leads to the copurification of components of the TIM23 complex [116, 123–127].

## VIII. Protein Transport Through Two Different Forms of the Presequence Translocase

The presequence translocase cooperates with the PAM complex for matrix protein transport. In contrast, precursor proteins with a hydrophobic sorting sequence are inserted into the inner membrane by the presequence

translocase in a manner independent of PAM. Analyses have now shown that indeed the presequence translocase exists in two distinct forms (1) a PAM-bound form ($TIM23^{PAM}$) and (2) a sorting competent but PAM-free form ($TIM23^{SORT}$) (Figure 15.2) [13, 58]. Thus, the molecular composition of the translocase adapts to the requirements of the translocating preprecursor, a process that must be highly regulated. Insight into how this process is regulated came from the identification of Tim21. Biochemical analyses show that Tim23, Tim17, and Tim50 are components of both $TIM23^{PAM}$ and $TIM23^{SORT}$ complexes [13, 45–48, 51, 58, 81]. In contrast, Tim21 can only be detected in the sorting competent $TIM23^{SORT}$ complex to which PAM is not bound [13, 51]. Thus, Tim21 is a selective constituent of the sorting competent form of the presequence translocase and its presence excludes binding of the PAM complex to the presequence translocase [13]. In mitochondria, $TIM23^{SORT}$ and $TIM23^{PAM}$ appear to coexist in equilibrium. This equilibrium can be influenced by changing the protein levels of Tim21 within mitochondria or by dissociation of the PAM complex from the TIM23 complex. Deletion of Tim21 promotes recruitment of the PAM complex to the presequence translocase and thus the formation of $TIM23^{PAM}$. In contrast, raising the mitochondrial levels of Tim21 shifts the equilibrium toward $TIM23^{SORT}$ and, in agreement with the concomitant loss of PAM, these mitochondria show selective defects in the transport of matrix proteins [13]. Based on these observations, a model of the dynamic changes of the presequence translocase between $TIM23^{SORT}$ and $TIM23^{PAM}$ was proposed in which Tim21 acts as a molecular switch that regulates the recruitment of the PAM complex or the formation of the sorting complex dependent on the requirements of the substrate (Figure 15.2).

## Acknowledgments

We thank Drs. W. Voos and D. Stojanovski for critical reading of the chapter and helpful discussions. Work of the author's laboratory is supported by the Deutsche Forschungsgemeinschaft, the Sonderforschungsbereich 388.

## References

1. Neupert, W. (1997). Protein import into mitochondria. *Annu Rev Biochem* 66:863–917.
2. Rehling, P., Brandner, K., and Pfanner, N. (2004). Mitochondrial import and the twin-pore translocase. *Nat Rev Mol Cell Biol* 5:519–530.
3. Koehler, C.M. (2004). New developments in mitochondrial assembly. *Annu Rev Cell Dev Biol* 20:309–335.

4. Endo, T., Yamamoto, H., and Esaki, M. (2003). Functional cooperation and separation of translocators in protein import into mitochondria, the double-membrane bounded organelles. *J Cell Sci* 116:3259–3267.
5. Jensen, R.E., and Dunn, C.D. (2002). Protein import into and across the mitochondrial inner membrane: role of the TIM23 and TIM22 translocons. *Biochim Biophys Acta* 1592:25–34.
6. Beddoe, T., and Lithgow, T. (2002). Delivery of nascent polypeptides to the mitochondrial surface. *Biochim Biophys Acta* 1592:35–39.
7. Mihara, K., and Omura, T. (1996). Cytoplasmic chaperones in precursor targeting to mitochondria: the role of MSF and hsp 70. *Trends Cell Biol* 6:104–108.
8. Pfanner, N., and Geissler, A. (2001). Versatility of the mitochondrial protein import machinery. *Nat Rev Mol Cell Biol* 2:339–349.
9. Perry, A.J., and Lithgow, T. (2005). Protein targeting: entropy, energetics and modular machines. *Curr Biol* 15:R423–R425.
10. Matouschek, A., Pfanner, N., and Voos, W. (2000). Protein unfolding by mitochondria: the Hsp70 import motor. *EMBO Rep* 1:404–410.
11. Pfanner, N., and Truscott, K.N. (2002). Powering mitochondrial protein import. *Nat Struct Biol* 9:234–236.
12. Neupert, W., and Brunner, M. (2002). The protein import motor of mitochondria. *Nat Rev Mol Cell Biol* 3:555–565.
13. Chacinska, A., Lind, M., Frazier, A.E., Dudek, J., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H.E., Truscott, K.N., Guiard, B., Pfanner, N., and Rehling, P. (2005). Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* 120:817–829.
14. Gakh, O., Cavadini, P., and Isaya, G. (2002). Mitochondrial processing peptidases. *Biochim Biophys Acta* 1592:63–77.
15. Voos, W., and Röttgers, K. (2002). Molecular chaperones as essential mediators of mitochondrial biogenesis. *Biochim Biophys Acta* 1592:51–62.
16. Schnell, D.J., and Hebert, D.N. (2003). Protein translocons. Multifunctional mediators of protein translocation across membranes. *Cell* 112:491–505.
17. Wickner, W., and Schekman, R. (2005). Protein translocation across biological membranes. *Science* 310:1452–1456.
18. Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schönfisch, B., Perschil, I., Chacinska, A., Guiard, B., Rehling, P., Pfanner, N., *et al.* (2003). The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc Natl Acad Sci USA* 100:13207–13212.
19. von Heijne, G. (1986). Mitochondrial targeting sequences may form amphiphilic helices. *EMBO J* 5:1335–1342.
20. Endo, T., and Kohda, D. (2002). Functions of outer membrane receptors in mitochondrial protein import. *Biochim Biophys Acta* 1592:3–14.
21. Pfanner, N. (2000). Protein sorting: recognizing mitochondrial presequences. *Curr Biol* 10:R412–R415.
22. Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nisnikawa, S.I., Endo, T., and Kohda, D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* 100:551–560.
23. Brix, J., Rüdiger, S., Bukau, B., Schneider-Mergener, J., and Pfanner, N. (1999). Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, Tom70 in a presequence-carrying and a non-cleavable protein. *J Biol Chem* 274:16522–16530.
24. Schatz, G. (1997). Just follow the acid chain. *Nature* 388:121–122.

25. Komiya, T., Rospert, S., Koehler, M., Looser, R., Schatz, G., and Mihara, K. (1998). Interaction of mitochondrial targeting signals with acidic receptor domains along the protein import pathway: evidence for the 'acid chain' hypothesis. *EMBO J* 17:3886–3898.
26. Bauer, M.F., Sirrenberg, C., Neupert, W., and Brunner, M. (1996). Role of Tim23 as voltage sensor and presequence receptor in protein import into mitochondria. *Cell* 87:33–41.
27. Truscott, K.N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A.J., Rassow, J., Pfanner, N., and Wagner, R. (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat Struct Biol* 8:1074–1082.
28. Pfanner, N., and Neupert, W. (1985). Transport of proteins into mitochondria: a potassium diffusion potential is able to drive the import of ADP/ATP carrier. *EMBO J* 4:2819–2825.
29. Martin, J., Mahlke, K., and Pfanner, N. (1991). Role of an energized inner membrane in mitochondrial protein import: $\Delta\psi$ drives the movement of presequences. *J Biol Chem* 266:18051–18057.
30. Endo, T., Mitsui, S., and Roise, D. (1995). Mitochondrial presequences can induce aggregation of unfolded proteins. *FEBS Lett* 359:93–96.
31. Luciano, P., and Geli, V. (1996). The mitochondrial processing peptidase: function and specificity. *Experientia* 52:1077–1082.
32. Nunnari, J., Fox, T.D., and Walter, P. (1993). A mitochondrial protease with two catalytic subunits of nonoverlapping specificities. *Science* 262:1997–2004.
33. Rospert, S., Junne, T., Glick, B.S., and Schatz, G. (1993). Cloning and disruption of the gene encoding yeast mitochondrial chaperonin 10, the homolog of *E. coli* groES. *FEBS Lett* 13:358–360.
34. Frazier, A.E., Dudek, J., Guiard, B., Voos, W., Li, Y., Lind, M., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H.E., Bilanchone, V., Cumsky, M.G., *et al.* (2004). Pam16 has an essential role in the mitochondrial protein import motor. *Nat Struct Mol Biol* 11:226–233.
35. Kozany, C., Mokranjac, D., Sichting, M., Neupert, W., and Hell, K. (2004). The J-domain related co-chaperone Tim16 is a constituent of the mitochondrial TIM23 preprotein translocase. *Nat Struct Mol Biol* 11:234–241.
36. Herrmann, J.M., and Hell, K. (2005). Chopped, trapped or tacked–protein translocation into the IMS of mitochondria. *Trends Biochem Sci* 30:205–211.
37. Dekker, P.J., Keil, P., Rassow, J., Maarse, A.C., Pfanner, N., and Meijer, M. (1993). Identification of MIM23, a putative component of the protein import machinery of the mitochondrial inner membrane. *FEBS Lett* 330:66–70.
38. Emtage, J.L.T., and Jensen, R.E. (1993). *MAS6* encodes an essential inner membrane component of the yeast mitochondrial protein import pathway. *J Biol Chem* 122:1003–1012.
39. Maarse, A.C., Blom, J., Keil, P., Pfanner, N., and Meijer, M. (1994). Identification of the essential yeast protein MIM17, an integral mitochondrial inner membrane protein involved in protein import. *FEBS Lett* 349:215–221.
40. Ryan, K.R., Menold, M.M., Garrett, S., and Jensen, R.E. (1994). SMS1, a high-copy suppressor of the yeast mas6 mutant, encodes an essential inner membrane protein required for mitochondrial protein import. *Mol Biol Cell* 5:529–538.
41. Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M., and Soll, J. (1999). The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J Mol Biol* 286:105–120.

42. Blom, J., Dekker, P.J., and Meijer, M. (1995). Functional and physical interactions of components of the yeast mitochondrial inner-membrane import machinery (MIM). *Eur J Biochem* 232:309–314.
43. Moro, F., Sirrenberg, C., Schneider, H.-C., Neupert, W., and Brunner, M. (1999). The TIM17.23 preprotein translocase of mitochondria: composition and function in protein transport into the matrix. *EMBO J* 18:3667–3675.
44. Ryan, K.R., Leung, R.S., and Jensen, R.E. (1998). Characterization of the mitochondrial inner membrane translocase complex: the Tim23p hydrophobic domain interacts with Tim17p but not with other Tim23p molecules. *Mol Cell Biol* 18:178–187.
45. Dekker, P.J., Martin, F., Maarse, A.C., Bömer, U., Müller, H., Guiard, B., Meijer, M., Rassow, J., and Pfanner, N. (1997). The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. *EMBO J* 16:5408–5419.
46. Geissler, A., Chacinska, A., Truscott, K.N., Wiedemann, N., Brandner, K., Sickmann, A., Meyer, H.E., Meisinger, C., Pfanner, N., and Rehling, P. (2002). The mitochondrial presequence translocase: an essential role of Tim50 in directing preproteins to the import channel. *Cell* 111:507–518.
47. Yamamoto, H., Esaki, M., Kanamori, T., Tamura, Y., Nishikawa, S., and Endo, T. (2002). Tim50 is a subunit of the TIM23 complex that links protein translocation across the outer and inner mitochondrial membranes. *Cell* 111:519–528.
48. Mokranjac, D., Paschen, S.A., Kozany, C., Prokisch, H., Hoppins, S.C., Nargang, F.E., Neupert, W., and Hell, K. (2003). Tim50, a novel component of the TIM23 preprotein translocase of mitochondria. *EMBO J* 22:816–825.
49. Kübrich, M., Keil, P., Rassow, J., Dekker, P.J.T., Blom, J., Meijer, M., and Pfanner, N. (1994). The polytopic mitochondrial inner membrane proteins MIM17 and MIM23 operate at the same preprotein import site. *FEBS Lett* 349:222–228.
50. Ryan, K.R., and Jensen, R.E. (1993). Mas6p can be cross-linked to an arrested precursor and interacts with other proteins during mitochondrial protein import. *J Biol Chem* 268:23743–23746.
51. Mokranjac, D., Popov-Celeketic, D., Hell, K., and Neupert, W. (2005). Role of Tim21 in mitochondrial translocation contact sites. *J Biol Chem* 280:23437–23440.
52. Lohret, T., Jensen, R.E., and Kinnally, K.W. (1997). Tim23, a protein import component of the mitochondrial inner membrane, is required for normal activity of the multiple conductance channel, MCC. *J Cell Biol* 137:377–386.
53. Schwartz, M.P., and Matouschek, A. (1999). The dimensions of the protein import channels in the outer and inner mitochondrial membranes. *Proc Natl Acad Sci USA* 96:13086–13090.
54. Kovermann, P., Truscott, K.N., Guiard, B., Rehling, P., Sepuri, N.B., Müller, H., Jensen, R.E., Wagner, R., and Pfanner, N. (2002). Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. *Mol Cell* 9:363–373.
55. Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H.E., Kühlbrandt, W., Wagner, R., Truscott, K.N., and Pfanner, N. (2003). Protein insertion into the mitochondria inner membrane by a twin-pore translocase. *Science* 299:1747–1751.
56. Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. *Nature* 395:516–521.
57. Athing, U., Thun, C., Hegerl, R., Typke, D., Nargang, F.E., Neupert, W., and Nussberger, S. (1999). The TOM core complex: the general protein import pore of the outer membrane of mitochondria. *J Cell Biol* 147:959–968.

58. Meinecke, M., Wagner, R., Kovermann, P., Guiard, B., Mick, D.U., Hutu, D.P., Voos, W., Truscott, K.N., Chacinska, A., Pfanner, P., and Rehling, P. (2006). Tim50 maintains the permeability barrier of the mitochondrial inner membrane. *Science* 5779:1523–1526.
59. Meier, S., Neupert, W., and Herrmann, J.M. (2005). Conserved N-terminal negative charges in the Tim17 subunit of the TIM23 translocase play a critical role in the import of preproteins into mitochondria. *J Biol Chem* 280:7777–7785.
60. Schleyer, M., Schmidt, B., and Neupert, W. (1982). Requirement of a membrane potential for the posttranslational transfer of proteins into mitochondria. *Eur J Biochem* 125:109–116.
61. Pfanner, N., Tropschug, M., and Neupert, W. (1987). Mitochondrial protein import: nucleoside triphosphates are involved in conferring import-competence to precursors. *Cell* 49:815–823.
62. Wachter, C., Schatz, G., and Glick, B.S. (1994). Protein import into mitochondria: the requirement for external ATP is precursor-specific whereas intramitochondrial ATP is universally needed for translocation into the matrix. *Mol Biol Cell* 5:465–474.
63. Hwang, S.T., and Schatz, G. (1989). Translocation of proteins across the mitochondrial inner membrane, but not into the outer membrane, requires nucleoside triphosphates in the matrix. *Proc Natl Acad Sci USA* 86:8432–8436.
64. Stuart, R.A., Gruhler, A., van der Klei, I., Guiard, B., Koll, H., and Neupert, W. (1994). The requirement of matrix ATP for the import of precursor proteins into the mitochondrial matrix and intermembrane space. *Eur J Biochem* 220:9–18.
65. Kang, P., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A., and Pfanner, N. (1990). Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature* 348:137–143.
66. Bukau, B., and Horwich, A.L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* 92:351–366.
67. Young, J.C., Agashe, V.R., Siegers, K., and Hartl, F.U. (2004). Pathways of chaperone-mediated protein folding in the cytosol. *Nat Rev Mol Cell Biol* 5:781–791.
68. Hartl, F.U., and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295:1852–1858.
69. Mayer, M.P., Brehmer, D., Gässler, C.S., and Bukau, B. (2002). Hsp70 chaperone machines. *Adv Protein Chem* 59:1–44.
70. Johnson, J.L., and Craig, E.A. (1997). Protein folding *in vivo*: unraveling complex pathways. *Cell* 90:201–204.
71. Kelley, W.L. (1998). The J-domain family and the recruitment of chaperone power. *Trends Biochem Sci* 23:222–227.
72. Kelley, W.L. (1999). Molecular chaperones: how J domains turn on Hsp70s. *Curr Biol* 9: R305–R308.
73. Walsh, P., Bursac, D., Law, Y.C., Cyr, D., and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep* 5:567–571.
74. Gambill, B.D., Voos, W., Kang, P.J., Miao, B., Langer, T., Craig, E.A., and Pfanner, N. (1993). A dual role for mitochondrial heat shock protein 70 in membrane translocation of preproteins. *J Cell Biol* 123:109–117.
75. Oka, T., and Mihara, K. (2005). A railroad switch in mitochondrial protein import. *Mol Cell* 18:145–146.
76. Maarse, A.C., Blom, J., Grivell, L.A., and Meijer, M. (1992). MPI1, an essential gene encoding a mitochondrial membrane protein, is possibly involved in protein import into yeast mitochondria. *EMBO J* 11:3619–3628.
77. Scherer, P.E., Manning-Krieg, U.C., Jenö, P., Schatz, G., and Horst, M. (1992). Identification of a 45-kDa protein at the protein import site of the yeast mitochondrial inner membrane. *Proc Natl Acad Sci USA* 89:11930–11934.

78. Blom, J., Kübrich, M., Rassow, J., Voos, W., Dekker, P.J., Maarse, A.C., Meijer, M., and Pfanner, N. (1993). The essential yeast protein MIM44 (encoded by MPI1) is involved in an early step of preprotein translocation across the mitochondrial inner membrane. *Mol Cell Biol* 13:7364–7371.
79. Horst, M., Jenö, P., Kronidou, N.G., Bolliger, L., Oppliger, W., Scherer, P., Manning-Krieg, U., Jascur, T., and Schatz, G. (1993). Protein import into yeast mitochondria: the inner membrane import site protein ISP45 is the *MPI1* gene product. *EMBO J* 12:3035–3041.
80. Merlin, A., Voos, W., Maarse, A.C., Meijer, M., Pfanner, N., and Rassow, J. (1999). The J-related segment of Tim44 is essential for cell viability: a mutant Tim44 remains in the mitochondrial import site, but inefficiently recruits mtHsp70 and impairs protein translocation. *J Cell Biol* 145:961–972.
81. Berthold, J., Bauer, M.F., Schneider, H.C., Klaus, C., Dietmeier, K., Neupert, W., and Brunner, M. (1995). The MIM complex mediates preprotein translocation across the mitochondrial inner membrane and couples it to the mt-Hsp70/ATP driving system. *Cell* 81:1085–1093.
82. Weiss, C., Oppliger, W., Vergeres, G., Demel, R., Jenö, P., Horst, M., de Kruijff, B., Schatz, G., and Azem, A. (1999). Domain structure and lipid interaction of recombinant yeast Tim44. *Proc Natl Acad Sci USA* 96:8890–8894.
83. Schneider, H.C., Berthold, J., Bauer, M.F., Dietmeier, K., Guiard, B., Brunner, M., and Neupert, W. (1994). Mitochondrial Hsp70/MIM44 complex facilitates protein import. *Nature* 371:768–774.
84. Kronidou, N.G., Oppliger, W., Bolliger, L., Hannavy, K., Glick, B.S., Schatz, G., and Horst, M. (1994). Dynamic interaction between Isp45 and mitochondrial hsp70 in the protein import system of the yeast mitochondrial inner membrane. *Proc Natl Acad Sci USA* 91:12818–12822.
85. Rassow, J., Maarse, A.C., Krainer, E., Kübrich, M., Müller, H., Meijer, M., Craig, E.A., and Pfanner, N. (1994). Mitochondrial protein import: biochemical and genetic evidence for interaction of matrix hsp70 and the inner membrane protein MIM44. *J Cell Biol* 127:1547–1556.
86. Krimmer, T., Rassow, J., Kunau, W.-H., Voos, W., and Pfanner, N. (2000). Mitochondrial protein import motor: the ATPase domain of matrix Hsp70 is crucial for binding to Tim44, while the peptide binding domain and the carboxy-terminal segment play a stimulatory role. *Mol Cell Biol* 20:5879–5887.
87. Strub, A., Lim, J.H., Pfanner, N., and Voos, W. (2000). The mitochondrial protein import motor. *Biol Chem* 381:943–949.
88. Strub, A., Röttgers, K., and Voos, W. (2002). The Hsp70 peptide-binding domain determines the interaction of the ATPase domain with Tim44 in mitochondria. *EMBO J* 21:2626–2635.
89. Bömer, U., Meijer, M., Maarse, A.C., Hönlinger, A., Dekker, P.J.T., Pfanner, N., and Rassow, J. (1997). Multiple interactions of components mediating preprotein translocation across the inner mitochondrial membrane. *EMBO J* 16:2205–2216.
90. Voos, W., Gambill, B.D., Laloraya, S., Ang, D., Craig, E.A., and Pfanner, N. (1994). Mitochondrial GrpE is present in a complex with hsp70 and preproteins in transit across membranes. *Mol Cell Biol* 14:6627–6634.
91. Ikeda, E., Yoshida, S., Mitsuzawa, H., Uno, I., and Toh-e, A. (1994). *YGE1* is a yeast homologue of *Escherichia coli* grpE and is required for maintenance of mitochondrial functions. *FEBS Lett* 339:265–268.
92. Laloraya, S., Gambill, B.D., and Craig, E.A. (1994). A role for a eukaryotic GrpE-related protein, Mge1p, in protein translocation. *Proc Natl Acad Sci USA* 91:6481–6485.

93. von Ahsen, O., Voos, W., Henninger, H., and Pfanner, N. (1995). The mitochondrial protein import machinery. Role of ATP in dissociation of the Hsp70.Mim44 complex. *J Biol Chem* 270:29848–29853.
94. Truscott, K.N., Voos, W., Frazier, A.E., Lind, M., Li, Y., Geissler, A., Dudek, J., Müller, H., Sickmann, A., Meyer, H.E., Meisinger, C., Guiard, B., *et al.* (2003). A J-protein is an essential subunit of the presequence translocase-associated protein import motor of mitochondria. *J Cell Biol* 163:707–713.
95. Mokranjac, D., Sichting, M., Neupert, W., and Hell, K. (2003). Tim14, a novel key component of the import motor of the TIM23 protein translocase of mitochondria. *EMBO J* 22:4945–4956.
96. D'Silva, P.D., Schilke, B., Walter, W., Andrew, A., and Craig, E.A. (2003). J protein cochaperone of the mitochondrial inner membrane required for protein import into the mitochondrial matrix. *Proc Natl Acad Sci USA* 100:13839–13844.
97. Voos, W., Gambill, B.D., Guiard, B., Pfanner, N., and Craig, E.A. (1993). Presequence and mature part of preproteins strongly influence the dependence of mitochondrial protein import on heat shock protein 70 in the matrix. *J Cell Biol* 123:119–126.
98. Li, Y., Dudek, J., Guiard, B., Pfanner, N., Rehling, P., and Voos, W. (2004). The presequence translocase-associated protein import motor of mitochondria: Pam16 functions in an antagonistic manner to Pam18. *J Biol Chem* 279:38047–38054.
99. D'Silva, P.R., Schilke, B., Walter, W., and Craig, E.A. (2005). Role of Pam16's degenerate J domain in protein import across the mitochondrial inner membrane. *Proc Natl Acad Sci USA* 102:12419–12424.
100. Mokranjac, D., Sichting, M., Popov-Celeketic, D., Berg, A., Hell, K., and Neupert, W. (2005). The import motor of the yeast mitochondrial TIM23 preprotein translocase contains two different J proteins, Tim14 and Mdj2. *J Biol Chem* 280:31608–31614.
101. van der Laan, M., Chacinska, A., Lind, M., Perschil, I., Sickmann, A., Meyer, H.E., Guiard, B., Meisinger, C., Pfanner, N., and Rehling, P. (2005). Pam17 is required for architecture and translocation activity of the mitochondrial protein import motor. *Mol Cell Biol* 25:7449–7458.
102. Rassow, J., Guiard, B., Wienhues, U., Herzog, V., Hartl, F.U., and Neupert, W. (1989). Translocation arrest by reversible folding of a precursor protein imported into mitochondria. A means to quantitate translocation contact sites. *J Cell Biol* 109:1421–1428.
103. Asai, T., Takahashi, T., Esaki, M., Nishikawa, S., Ohtsuka, K., Nakai, M., and Endo, T. (2004). Reinvestigation of the requirement of cytosolic ATP for mitochondrial protein import. *J Biol Chem* 279:19464–19470.
104. Matouschek, A., Azem, A., Ratliff, K., Glick, B.S., Schmid, K., and Schatz, G. (1997). Active unfolding of precursor proteins during mitochondrial protein import. *EMBO J* 16:6727–6736.
105. Huang, S., Ratliff, K.S., and Matouschek, A. (2002). Protein unfolding by the mitochondrial membrane potential. *Nat Struct Biol* 9:301–307.
106. Huang, S., Ratliff, K.S., Schwartz, M.P., Spenner, J.M., and Matouschek, A. (1999). Mitochondria unfold precursor proteins by unraveling them from their N-termini. *EMBO J* 6:1132–1138.
107. Geissler, A., Rassow, J., Pfanner, N., and Voos, W. (2001). Mitochondrial import driving forces: enhanced trapping by matrix Hsp70 stimulates translocation and reduces the membrane potential dependence of loosely folded preproteins. *Mol Cell Biol* 21:7097–7104.
108. Voisine, C., Craig, E.A., Zufall, N., von Ahsen, O., Pfanner, N., and Voos, W. (1999). The protein import motor of mitochondria: unfolding and trapping of preproteins are distinct and separable functions of matrix Hsp70. *Cell* 97:565–574.

109. Moro, F., Okamoto, K., Donzeau, M., Neupert, W., and Brunner, M. (2002). Mitochondrial protein import: molecular basis of the ATP-dependent interaction of MtHsp70 with Tim44. *J Biol Chem* 277:6874–6880.
110. Liu, Q., D'Silva, P., Walter, W., Marszalek, J., and Craig, E.A. (2003). Regulated cycling of mitochondrial Hsp70 at the protein import channel. *Science* 300:139–141.
111. Mayer, M.P. (2004). Timing the catch. *Nat Struct Mol Biol* 11:6–8.
112. D'Silva, P., Liu, Q., Walter, W., and Craig, E.A. (2004). Regulated interactions of mtHsp70 with Tim44 at the translocon in the mitochondrial inner membrane. *Nat Struct Mol Biol* 11:1084–1091.
113. Schwaiger, M., Herzog, V., and Neupert, W. (1987). Characterization of translocation contact sites involved in the import of mitochondrial proteins. *J Cell Biol* 105:235–246.
114. Rassow, J., Hartl, F.U., Guiard, B., Pfanner, N., and Neupert, W. (1990). Polypeptides traverse the mitochondrial envelope in an extended state. *FEBS Lett* 275:190–194.
115. Donzeau, M., Káldi, K., Adam, A., Paschen, S., Wanner, G., Guiard, B., Bauer, M.F., Neupert, W., and Brunner, M. (2000). Tim23 links the inner and outer mitochondrial membranes. *Cell* 101:401–412.
116. Chacinska, A., Rehling, P., Guiard, B., Frazier, A.E., Schulze-Specking, A., Pfanner, N., Voos, W., and Meisinger, C. (2003). Mitochondrial translocation contact sites: separation of dynamic and stabilizing elements in formation of a TOM-TIM-preprotein supercomplex. *EMBO J* 22:5370–5381.
117. Frazier, A.E., Chacinska, A., Truscott, K.N., Guiard, B., Pfanner, N., and Rehling, P. (2003). Mitochondria use different mechanisms for transport of multispanning membrane proteins through the intermembrane space. *Mol Cell Biol* 23:7818–7828.
118. Moczko, M., Bömer, U., Kübrich, M., Zufall, N., Höhnlinger, A., and Pfanner, N. (1997). The intermembrane space domain of mitochondrial Tom22 functions as a *trans* binding site for preproteins with N-terminal targeting sequences. *Mol Cell Biol* 17:6574–6584.
119. Bolliger, L., Junne, T., Schatz, G., and Lithgow, T. (1995). Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria. *EMBO J* 14:6318–6326.
120. Kanamori, T., Nishikawa, S., Nakai, M., Shin, I., Schultz, P.G., and Endo, T. (1999). Uncoupling of transfer of the presequence and unfolding of the mature domain in precursor translocation across the mitochondrial outer membrane. *Proc Natl Acad Sci USA* 96:3634–3639.
121. Ohba, M., and Schatz, G. (1987). Disruption of the outer membrane restores protein import to trypsin-treated yeast mitochondria. *EMBO J* 6:2117–2122.
122. Hwang, S., Jascur, T., Vestweber, D., Pon, L., and Schatz, G. (1989). Disrupted yeast mitochondria can import precursor proteins directly through their inner membrane. *J Cell Biol* 109:487–493.
123. Horst, M., Hilfiker-Rothenfluh, S., Oppliger, W., and Schatz, G. (1995). Dynamic interaction of the protein translocation systems in the inner and outer membranes of yeast mitochondria. *EMBO J* 14:2293–2297.
124. Dekker, P.J., Martin, F., Maarse, A.C., Bömer, U., Müller, H., Guiard, B., Meijer, M., Rassow, J., and Pfanner, N. (1997). The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. *EMBO J* 16:5408–5419.
125. Schülke, N., Sepuri, N.B., Gordon, D.M., Saxena, S., Dancis, A., and Pain, D. (1999). A multisubunit complex of outer and inner mitochondrial membrane protein translocases stabilized *in vivo* by translocation intermediates. *J Biol Chem* 274:22847–22854.

126. Schülke, N., Sepuri, N.B., and Pain, D. (1997). *In vivo* zippering of inner and outer mitochondrial membranes by a stable translocation intermediate. *Proc Natl Acad Sci USA* 94:7314–7319.
127. Wienhues, U., Becker, K., Schleyer, M., Guiard, B., Tropschug, M., Horwich, A.L., Pfanner, N., and Neupert, W. (1991). Protein folding causes an arrest of preprotein transloction into mitochondria *in vivo*. *J Cell Biol* 115:1601–1609.

This page intentionally left blank

# Part IV

## *Crossing Chloroplast Membranes*

This page intentionally left blank

# 16

# *The Toc Machinery of the Protein Import Apparatus of Chloroplasts*

CALEB ROUNDS[a,b] • FEI WANG[a,c] • DANNY J. SCHNELL[a,b,c]

[a]*Department of Biochemistry and Molecular Biology*
[b]*Program in Plant Biology*
[c]*Program in Molecular and Cellular Biology*
*University of Massachusetts*
*Amherst, MA 01003, USA*

## I. Introduction

Chloroplasts contain an elaborate architecture that has evolved to optimize the reactions of photosynthesis, regulate the partitioning of photosynthate metabolites, and participate in essential steps of amino acid, lipid, and secondary metabolism [1]. The organelle contains at least six suborganellar compartments that are defined by three independent membrane systems, the outer and inner envelope membranes and the thylakoid membrane. Chloroplast biogenesis is dependent on contributions from both the organellar and nuclear genomes in plant cells. An estimated 2500–3500 nuclear genes encode chloroplast proteins in *Arabidopsis thaliana* [2–4]. This number corresponds to 10–15% of the total genes in the nuclear genome. The remaining ~120 genes are encoded in the organellar genome. The predominant role of the nucleus in chloroplast biogenesis requires that thousands of nucleus-encoded polypeptides be imported into the organelle from their site of synthesis in the cytoplasm and accurately assembled into functional protein complexes (Figure 16.1). The initial steps in the recognition and translocation of proteins into chloroplasts are mediated by translocons at the outer envelope membrane of chloroplasts (Toc complexes) [5].



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25016-4

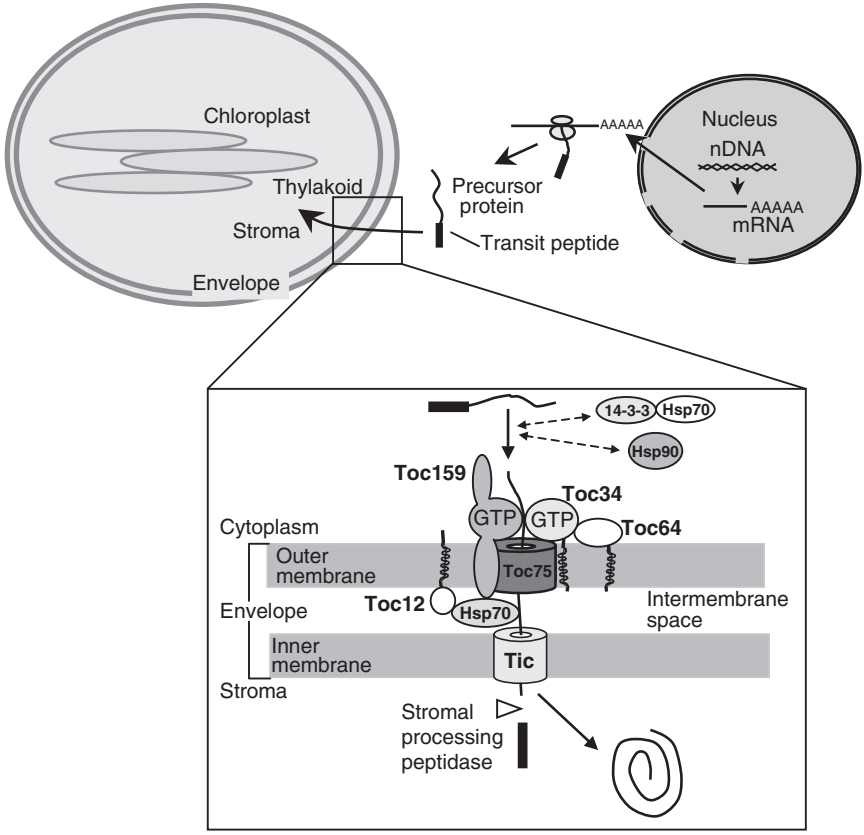



FIG. 16.1. Schematic representation of protein import into chloroplasts and the composition of Toc complexes. Nucleus-encoded chloroplast preproteins are imported into the organelle after the completion of synthesis on cytoplasmic ribosomes. The recognition and translocation of preproteins across the outer and inner envelope membranes is directed by their N-terminal, cleavable transit peptides. The initial docking of preproteins and their insertion across the outer membrane is mediated by the Toc translocon. The translocon is composed of three core components, Toc34, Toc159, and Toc75. Toc34 and Toc159 form a transit peptide receptor system that is regulated by GTP binding and hydrolysis. In addition, the receptors regulate the transfer of preproteins to Toc75, the major component of the Toc protein-conducting channel. Targeting of preproteins from the cytoplasm to the chloroplast surface might be assisted by cytoplasmic factors that include molecular chaperones. These complexes include a guidance complex containing 14-3-3 and Hsp70 proteins, and an Hsp90 that binds to Toc64 at the chloroplast surface. Once inserted across the outer membrane, the binding of an Hsp70 chaperone in the intermembrane space assists translocation through the channel. Toc12 forms a potential docking site for the Hsp70 at the inner face of the outer membrane.

These multisubunit Toc complexes physically associate with translocons at the inner envelope membrane of chloroplasts (Tic complexes) to provide a direct conduit from the cytoplasm to the chloroplast interior.

A general picture of Toc complex composition is now in hand (Figure 16.1). This view reveals the existence of a subset of components that are homologous to prokaryotic membrane transport components, suggesting an evolutionary process in which existing membrane transporters were adapted for the novel function of protein import during endosymbiosis [6]. The selectivity and directionality of the import process appears to have evolved via the acquisition of unique targeting and receptor components that assembled with the transporters to generate highly specific protein translocons. Most studies on the import process predicted the existence of a general translocon that provides a single portal for protein translocation across the outer membrane [7, 8]. Although this appears to be the case in simple photosynthetic eukaryotes, studies have revealed the existence of multiple, distinct Toc complexes in complex land plants [9–11]. This diversity of Toc complexes appears to constitute a multipathway system for the targeting of distinct subclasses of nucleus-encoded proteins into the organelle. These pathways appear to play a critical role in the development and maintenance of diverse plastid types within specific organs of multicellular plants.

This chapter will focus on the molecular mechanism of Toc complex function in chloroplast protein import. In addition to the structural, biochemical, and genetic data that serve as the basis for hypotheses on Toc function, we also will include a comparative analysis of the Toc complexes within *Arabidopsis* and between different plant species. These comparisons are particularly informative in understanding what constitutes the basic unit of Toc complex function and what components likely contribute regulatory or selective functions on the import apparatus.

## II. General Overview of Toc Complexes

The general principles that govern translocon structure and function at other organellar membranes also apply to Toc complexes (Figure 16.1) [12, 13]. The Toc translocon contains a set of surface-exposed receptors that confer selectivity on the import process by specifically recognizing the intrinsic targeting signals (transit peptides) of nucleus-encoded chloroplast preproteins in the cytoplasm [14–17]. Protein import is a posttranslational process that does not appear to directly couple protein synthesis with targeting. The preprotein receptors are linked to a protein-conducting channel that provides an aqueous pore for the translocation of protein

substrates across the membrane [16, 18, 19]. Membrane translocation requires an energetic driving force, and, in the case of the Toc translocon, this appears to be fulfilled by nucleoside triphosphate hydrolysis at the receptors or at molecular chaperones that associate with Toc complexes [20–23]. The chaperones also appear to assist in maintaining the translocation substrate in an import competent conformation [24–28].

Biochemical approaches based on *in vitro* import assays with isolated chloroplasts have led to the identification of all known Toc components. The bulk of this chapter will focus on a set of three core Toc components, Toc75, Toc34, and Toc159 (the numbers refer to their molecular masses in kilodaltons) [5]. These Toc proteins are defined as core components because genetic analysis in *Arabidopsis* demonstrates that they are essential for protein import, and therefore plant viability [29–32]. In addition, biochemical studies provide compelling evidence for their functions within the Toc translocon [14–17]. Toc34 and Toc159 are membrane-integrated GTPases that form the transit peptide receptor system at the translocon [16, 19]. The experimental evidence indicates that Toc75 is a β-barrel membrane protein that forms a major component of the protein-conducting channel [16, 19, 33]. A molecular chaperone of the Hsp70 family, Hsp70-IAP, associates with the Toc complex in the intermembrane space [28, 34, 35]. Hsp70-IAP70 is proposed to assist in import by providing a component of the translocation motor that drives preproteins across the Toc channel.

The three core Toc components form a complex in the chloroplast outer envelope membrane under steady state conditions. These complexes are 500–800 kDa in size, indicating that multiple copies of each core component are present within individual translocons [36]. Toc34 and Toc75 appear to be present in a 1:1 stoichiometry within these complexes, whereas Toc159 is present in substoichiometric amounts relative to Toc34 and Toc75 [36]. These isolated complexes can mediate protein translocation in reconstituted proteoliposomes, providing direct evidence that they are sufficient to support translocation across the outer membrane [37].

## III. Toc Receptors

Toc34 and Toc159 form a coordinated transit peptide receptor system at the chloroplast surface [16, 19]. Analysis of the Toc34 and Toc159 receptor families in *Arabidopsis* demonstrates that members of both classes of Toc GTPase are essential for chloroplast biogenesis and plant viability [30, 32, 38]. Therefore, the activities of both receptors are required for the import process. In the absence of energy, both Toc34 and Toc159 cross-link to bound chloroplast preproteins in isolated chloroplasts [16, 19, 20], indicating that

they participate together in the initial steps of preprotein recognition at the chloroplast surface. This interaction is transit peptide specific and readily reversible, and is believed to represent the initial, specific binding of the transit peptide at the Toc complex [19, 20, 39–41]. Nonhydrolyzable GTP analogues block import at this stage, indicating that GTP hydrolysis at the Toc GTPases promotes transfer of the preprotein into the protein-conducting channel [19, 23].

The bulk of both receptors face the cytoplasm with their C-terminal regions tethered to the membrane. Toc34 contains an ~30-kDa cytoplasmic GTPase domain that is anchored to the outer membrane with a short α-helical transmembrane segment (Figure 16.2A) [15, 17]. The Toc159 receptor is composed of a tripartite domain structure (Figure 16.2A). It contains an N-terminal A-domain (~70 kDa) that is characterized by tandem repeats of alternating acidic and hydrophobic residues [42–44]. The function of this domain has not been determined. The A-domain is followed by the GTPase or G-domain. The size and primary structure of the G-domain is similar to that of Toc34 [15, 17]. The C-terminal M-domain (~52 kDa) of Toc159 anchors the protein in the outer membrane. The M-domain is protected from external proteolysis and is not extracted from the outer membrane by aqueous perturbants, indicating that it is integrated into the lipid bilayer [15, 17, 42, 43]. Despite these biochemical indications of membrane integration, the M-domain lacks characteristic hydrophobic segments that could serve as transmembrane helices. It has been proposed that the M-domain might interface with the bilayer via β-strands or by exposing residues to the core of the bilayer as it assembles with the other components of the Toc complex.

The GTPase domains of the Toc receptors contain amino acid sequences that correspond to the five conserved motifs of the typical Ras-type GTPases (G1–G5) (Figure 16.2B). On the basis of this observation, it was assumed that the structure of the Toc GTPase domains would mimic the canonical GTPases [15, 17]. On the contrary, the X-ray crystal structure of the GTPase domain of Toc34 from pea revealed a reorganized set of motifs that result in a significantly distinct class of GTPases [45]. The G1 and G3 motifs that include the universal P-loop nucleotide-binding residues are conserved in Toc34. However, the conserved G2, G4, and G5 motifs of Ras and other GTPases are rearranged in Toc34 and no longer participate in nucleotide binding. Instead, these motifs have been substituted with new molecular interactions that form the nucleotide-binding site [45]. These structural changes suggest a novel mechanism of coupling nucleotide state with conformational changes in the Toc GTPases, a prediction that is reflected in some of the unique functional characteristics of the receptors (see below). Toc159 is predicted to have a similar G-domain structure as that observed in Toc34 based on the high degree of sequence similarity between these regions of the proteins [45].

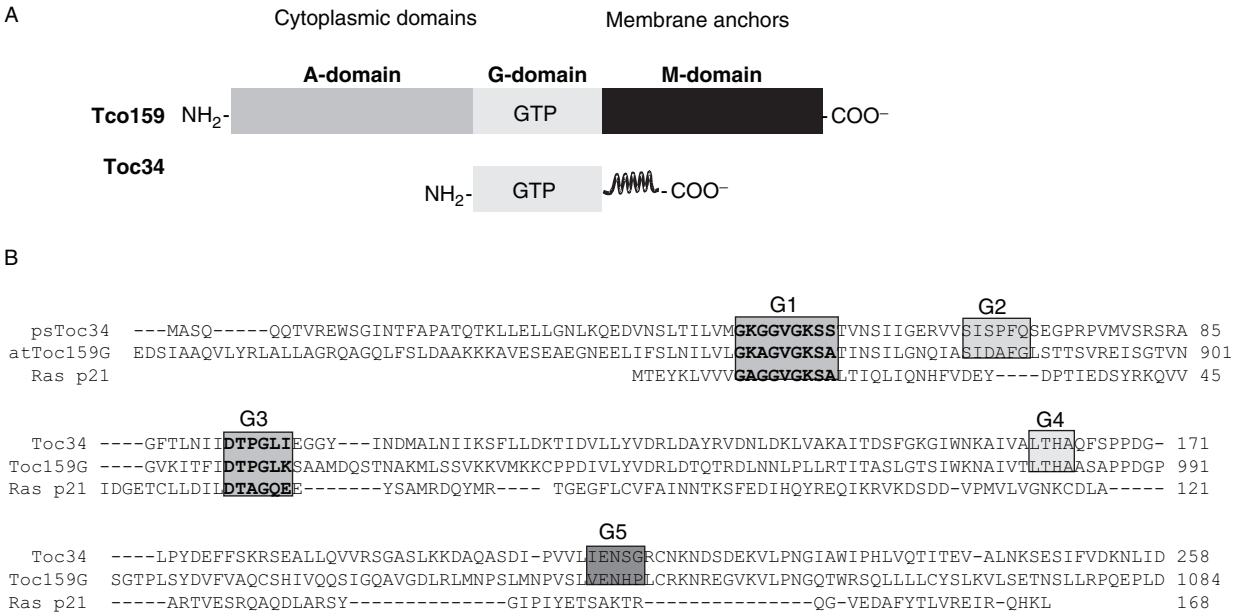



FIG. 16.2. Structural comparisons of the Toc GTPases. (A) Schematic representation of the domain structures of Toc159 and Toc34. Toc34 contains a cytoplasmic GTPase domain (G-domain) that is anchored to the outer membrane with a single α-helical transmembrane helix. Toc159 contains a central G-domain that is structurally homologous to that of Toc34. The G-domain is flanked by an N-terminal acidic domain (A-domain) of unknown function and a C-terminal M-domain that anchors the protein to the outer membrane. (B) Comparison of the primary structures of the G-domains of pea Toc34 (psToc34) and *Arabidopsis* Toc159 (atToc159) to Ras-p21. The G1–G5 motifs involved in GDP binding in the psToc34 crystal structure and their conserved sequences within atToc159 or Ras-p21 are highlighted by boxes. PsToc34 contains the consensus G1 (P-loop) and G2 motifs that are found in all regulatory GTPases, including Ras. The G2, G4, and G5 motifs of psToc34 are conserved in atToc159, but are distinct from those of Ras and other GTPases. This figure is adapted from Sun *et al*. [45].

The kinetics of Toc GTPase activity have been studied most extensively for atToc33, the *Arabidopsis* orthologue of Toc34. Similar to other regulatory GTPases, at Toc33 has a low intrinsic rate of hydrolysis ($k_{cat}$ = ~0.013–0.31 $min^{-1}$). Its affinities for GDP ($K_d$ = ~4.5 μM) and GTP ($K_d$ = ~2.6 μM) are similar [39, 46, 47]. These affinities are significantly lower than those of Ras ($K_d$ = ~0.0001–0.1 μM), but similar to those measured for SRP/SR ($K_d$ = 1–10 μM) [47]. Full-length Toc159 has not been purified, and therefore, its GTPase activity only has been measured in fragments consisting of its G-domain or G- and M-domains [48]. These studies suggest that its GTPase activity is approximately twofold higher than that of Toc34.

The crystal structure of Toc34 provided key information on the potential mechanism by which GTPase activity controls preprotein recognition by the Toc receptors. The Toc34 GTPase domain forms a dimer with a twofold axis of symmetry in the crystal [45]. The GTP-binding pockets of each monomer lie within the dimer interface, and each monomer extends residues into the nucleotide-binding site of the reciprocal monomer. Interestingly, polypeptide loops in the dimer form a cage around the bound GDP and appear to preclude nucleotide exchange. Furthermore, Glu73 of Toc34 lies adjacent to the β-phosphate of the bound GDP, and therefore this residue would need to relocate on GTP binding to accommodate the γ-phosphate. These structural data predict that the form of bound nucleotide (i.e., GTP versus GDP) strongly influence dimerization. Alternatively, changes in dimerization would be required to allow nucleotide exchange. Subsequent binding studies demonstrated that Toc34-Toc159 dimerization via interactions that involve their G-domains also occurs and is controlled by nucleotide state [46–50].

A variety of experimental approaches combined with the structural data have led to at least two possible roles of Toc GTPase activity. The first proposes that Toc34 and Toc159 form a GTP-regulated gate that controls the transition from preprotein binding at the chloroplast surface to membrane translocation during the import process (Figure 16.3A) [17, 51]. In this hypothesis, transit peptide binding to the receptors would stimulate GTPase activity and lead to changes in homo- or heterodimerization. This would constitute a molecular switch that opens the gate to the translocon and results in transfer of the preprotein into the protein-conducting channel. This model is supported by the observation that transit peptide binding is mediated, at least in part, by the G-domains of both receptors, and binding stimulates their GTPase activities, providing evidence for a direct interaction between preprotein recognition and GTPase activity [40, 46, 52]. Furthermore, studies have shown that Toc34-Toc159 dimerization is nucleotide dependent [48, 49]. This model is somewhat analogous to the role of GTP in regulating the association of the signal recognition particle (SRP) and its

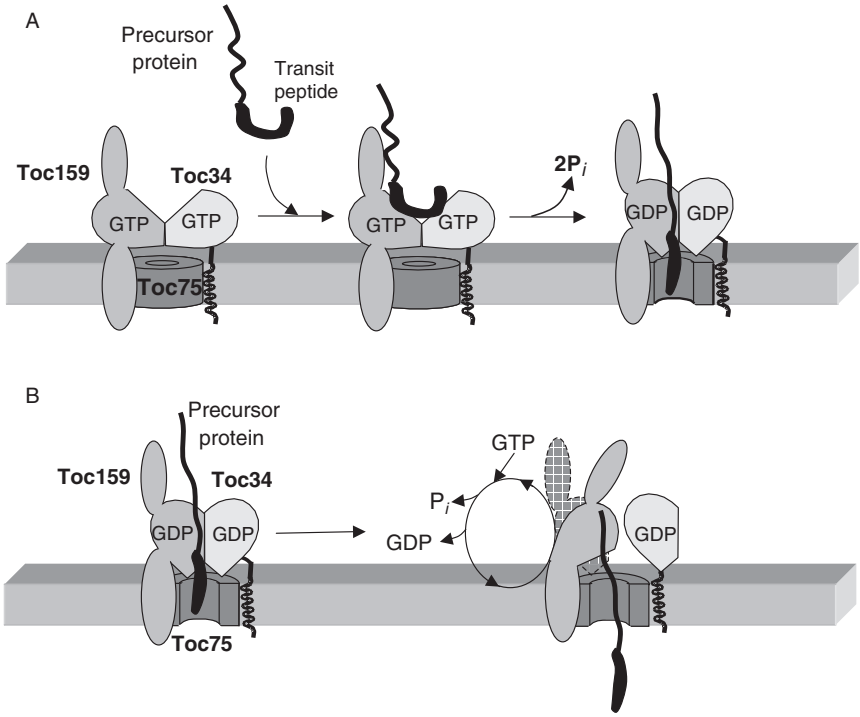



FIG. 16.3. Two proposed roles for the GTPase activities of Toc34 and Toc159 in preprotein recognition by the Toc translocon. (A) The Toc GTPases are proposed to regulate preprotein recognition by forming a GTP-regulated gate to the translocon. In this model, transit peptide binding at Toc159 and Toc34 stimulates the GTPase activities of the receptors. Hydrolysis is proposed to open the gate to the translocon by altering homo- and/or heterodimerization of the receptors. This, in turn, results in transfer of the preprotein into the translocon channel. (B) In addition to regulating preprotein recognition, the GTPase activity of Toc159 is proposed to provide the energy for preprotein translocation through the Toc channel. In this model, Toc159 acts as a motor that drives translocation through repetitive cycles of GTP-dependent preprotein binding and release.

receptor (SR) during the targeting of nascent chain ribosome complexes to the Sec61p translocon of the endoplasmic reticulum (ER) [53]. The interchange between homo- and heterodimeric states as a potential regulator of Toc complex function is very appealing, and future studies should demonstrate the importance of dimerization in the targeting reaction.

A second proposed function for Toc GTPase activity, in addition to translocon gating, is to provide the energetic driving force for translocation across the outer membrane (Figure 16.3B) [37]. In this model, Toc159 is proposed to function as a molecular motor that drives translocation via

repeated cycles of GTP binding and hydrolysis in addition to contributing to translocon gating. This activity is supported by the observation that Toc159 and Toc75 alone can mediate a minimum level of GTP-dependent preprotein insertion into reconstituted proteoliposomes [37]. Although multiple lines of evidence support a role for GTP hydrolysis in transferring the preprotein from the Toc GTPases to the translocon channel, several observations indicate that repeated GTP hydrolysis at Toc159 is not required for complete translocation. First, nonhydrolyzable GTP analogues do not block the translocation of preproteins across the outer membrane once they have initially inserted into the Toc translocon channel [23]. Translocation from this early intermediate step requires only ATP, implicating molecular chaperones as the driving force for import. Second, the GTPase activity of Toc159 does not appear to be an absolute requirement for import. Preprotein import occurs in chloroplasts in which the Toc159 A- and G-domains have been proteolytically removed [44]. Moreover, a Toc159 deletion construct, consisting solely of the M-domain, can partially complement a Toc159 null mutant in *Arabidopsis* [54]. A possible explanation to reconcile the results from the reconstitution studies with those in isolated chloroplast and *in vivo* is that Toc159 GTPase activity might contribute to dissociation of the preproteins from the receptor complex and transfer to the Toc channel, thereby facilitating insertion in the channel. In the reconstituted system, this activity might be sufficient to allow some transport across the membrane in the absence of Toc34. However, in intact Toc complexes, the activity of the Toc34 GTPase also appears to be important in the initiation of the translocation reaction [55].

## IV. The Toc Translocon Channel and Membrane Translocation

The function of Toc75 as the major component of the protein-conducting channel of the Toc translocon is supported both by biochemical and electrophysiological studies. In the presence of low concentrations of GTP and ATP, the bound preprotein is transferred from the Toc GTPase receptors into the channel of the translocon [20]. Protease protection studies indicate that the preprotein is inserted across the outer membrane and makes contact with the translocon at the inner membrane. At this stage, both the transit peptide and mature regions of the polypeptide form predominant cross-links to Toc75, indicating that it directly participates in membrane translocation. Direct evidence that Toc75 forms a membrane channel was provided by electrophysiological studies using recombinant Toc75 reconstituted into lipid bilayers [18]. These studies demonstrated that Toc75

forms a relatively unselective channel whose conductance is responsive to the binding of preproteins. The conductance measurements indicate that Toc75 can form a channel with a diameter of ~14 Å. This pore is large enough to accommodate an unfolded polypeptide chain [18].

The structural similarities of Toc75 to Tom40, the translocon channel of the mitochondrial outer membrane, and to bacterial outer membrane solute transporters also support the assignment of Toc75 as the translocon channel [56]. Although there is no apparent identity at the level of primary structure, Toc75 and Tom40 are both members of the β-barrel integral membrane protein family. Toc75 is deeply embedded in the outer membrane with up to 16 membrane-spanning β-strands.

Interestingly, Toc75 has been implicated in the insertion of membrane proteins into the Toc [11]. This process appears to be independent of the Toc GTPases, suggesting that Toc75 might function as a relatively nonselective polypeptide channel that assembles with different receptors to mediate the transport or insertion of a variety of proteins at the outer membrane. This proposal is consistent with the observation that the amounts of Toc75 exceed those of the Toc GTPases, and a significant portion of Toc75 is not found in complexes with the Toc GTPase receptors [57].

Although it is clear that Toc75 is an essential component of protein-conducting channels, covalent cross-linking and genetic studies suggest that Toc159 also participates in the translocation reaction [19, 20, 30]. The M-domain of Toc159 forms covalent cross-links with early import intermediates that are inserted across the membrane [20]. This observation suggests that the M-domain assists in the formation of the channel or interacts with the preprotein in the intermembrane space as it crosses the outer membrane. A role for the M-domain in preprotein translocation also is supported by the observation that overexpression of the M-domain alone can partially complement a Toc159 null mutant in *Arabidopsis* [54].

As mentioned previously, the Hsp70-IAP in the intermembrane space of the chloroplast envelope is proposed to bind to preproteins and provide the unidirectional driving force for translocation across the outer membrane [28, 34]. Although GTP hydrolysis at the Toc receptors appears to be required to transfer the preprotein into the channel, ATP hydrolysis in the intermembrane space is necessary for the formation of early intermediates that stably span the outer membrane [21, 23]. The Hsp70-IAP is proposed to use the ATP to bind and trap the polypeptide, thereby preventing it from slipping out of the translocon. The Hsp70-IAP likely also plays a role in preventing the folding or aggregation of preproteins in the intermembrane space prior to their insertion in the Tic translocon at the inner membrane. Translocation through the Toc translocon appears to require an unfolded substrate because preproteins with covalently

stabilized tertiary structures are not imported into chloroplasts [58–60]. A small outer membrane protein, Toc12, was discovered that contains a J-domain similar to those found in Hsp70 cochaperones in bacteria [61]. Toc12 is an excellent candidate for a protein that coordinates association of the Hsp70-IAP with the core Toc complex.

## V. Cytoplasmic Events

Several studies suggest that cytoplasmic factors aid in the delivery and translocation of preproteins at the Toc translocon (Figure 16.1). Two different studies implicate cytoplasmic Hsp70s in targeting, suggesting that this class of chaperone might play a role in preprotein targeting as has been shown in mitochondria and the ER. Preproteins cross-link with a cytoplasmic Hsp70 that associates with the chloroplast surface [62, 63]. Furthermore, an Hsp70 and a 14-3-3 protein were shown to associate with the precursor to the small subunit of rubisco (preSSU) in wheat germ extracts [63]. Fractions of preSSU containing the 14-3-3 and Hsp70 were shown to have a higher efficiency of import compared to those lacking the factors. Based on these and subsequent studies, it was proposed that the 14-3-3 protein and Hsp70 form a guidance complex that docks at Toc34 and delivers the preprotein to the Toc complex [64]. The association of guidance complex with the preprotein appeared to require phosphorylation of the preSSU transit peptide. Subsequent studies demonstrated that mutagenesis of the phosphorylation site on the preSSU transit peptide did not affect its targeting *in vivo* [65]. Furthermore, preSSU from a number of plant species lack phosphorylation sites. Therefore, it is possible that the guidance complex plays an accessory but nonessential role in targeting. The specific role of Hsp70 in import also is not clear because a stimulation of import by the addition of purified Hsp70 has not been demonstrated, and genetic studies have not implicated this chaperone in import [65].

*In vitro* binding studies have implicated a complex involving Hsp90 and an outer membrane protein, Toc64, in the targeting of a subset of preproteins to Toc complexes [64]. Toc64 is an integral outer membrane protein that is loosely associated with the core Toc components through a proposed interaction with Toc34 [64, 66]. Toc64 also appears to interact with a cytoplasmic Hsp90 [64]. This led to the proposal that the Hsp90 binds to preproteins in the cytoplasm and docks at Toc64 at the outer membrane. Delivery of the preprotein-Hsp90-Toc64 complex to the translocon is proposed to occur via a GTP-dependent interaction with Toc34. The potential role of Hsp90-Toc64 system appears to be specialized. To date, only a single

preprotein has been shown to interact with these factors. Furthermore, double null mutants of the two Toc64 isoforms in moss do not exhibit a detectable phenotype [67], and a Toc64 null mutant in *Arabidopsis* exhibits only a modest defect in the import of one preprotein, preOE33 [64]. It is possible that the ''guidance complex'' and/or the Hsp90/Toc64 systems represent scavenger processes or mechanisms of regulating the import process under certain physiological conditions (e.g., stress and fluctuations in import demand). This would explain why bypassing these systems does not have apparent effects under normal growth conditions. Additional *in vivo* experiments will be important in defining the specific roles of these components in the import process.

## VI. Toc Complex Evolution and Diversity

Chloroplasts are hypothesized to have arisen by the endosymbiotic assimilation into a nucleated cell of a photosynthetic bacterium similar to a cyanobacterium [68]. The evolution of the bacterium into a chloroplast resulted in the transfer of several thousand genes from the prokaryotic genome to the nucleus, and the concomitant development of a protein import system to compensate for the fact that the synthesis of proteins required for organelle biogenesis was displaced to the cytoplasm [4]. A protein import pathway comparable to those in chloroplasts is not known to exist in Gram-negative bacteria, suggesting that this trafficking pathway evolved anew in response to gene transfer [69]. Analysis of the core Toc components provides insight into the evolution of the import apparatus in response to endosymbiosis. Furthermore, the availability of EST and genome information from photosynthetic eukaryotes that closely resemble what are believed to be early stages in endosymbiosis provide information on the basic unit required for protein import.

The outer and inner chloroplast envelope membranes are evolutionarily related to the outer and cytoplasmic membrane of Gram-negative photosynthetic bacteria, respectively. Of the Toc components, only Toc75 has clear homologues in bacteria [6]. A closely related gene in cyanobacteria, SynToc75, is essential for viability [42, 43, 70]. The protein is located in the outer membrane of the bacterium and has demonstrated channel activity. However, the precise role of this protein is not known. Toc75 and SynToc75 are related to the Omp85 family of bacterial outer membrane proteins. Omp85 is implicated in bacterial outer membrane biogenesis by assisting in the integration of integral membrane proteins [71]. Other more distant relatives of the Omp85 family are known peptide transporters. This has

led to the hypothesis that Toc75 evolved from the preexisting bacterial peptide transporters to assume a new function in protein import.

The evolutionary origins of the Toc GTPases are unclear with no readily apparent evolutionary homologues in bacteria or other organisms outside of plants. The only similarities to proteins other than their orthologues in plants are other GTPases. Yet, the GTPase domains of the Toc receptors appear to have diverged sufficiently to prevent the establishment of evolutionary links. Perhaps the most interesting information on the evolution of Toc complexes comes from examining genomic data from a wide variety of photosynthetic origins that span the phylogenetic spectrum [72, 73]. Genomic analysis of simple photosynthetic eukaryotes, such as the red algae, indicates the presence of Toc75 and Toc34 orthologues [73, 74]. Proteins comparable to the Toc159 receptors are not present. Comparative protein import studies demonstrate that the basic mechanism of protein import is conserved between these simple organisms and vascular plants [72, 74, 75] Thus, it appears that Toc75 and Toc34 constitute the most basic unit of Toc complexes that can mediate protein import into chloroplasts.

The addition of the Toc159 receptor family to the Toc translocon appears to coincide with the evolution of more complex plant lineages, including multicellular plants that undergo cellular differentiation [72, 73]. In higher plants, cellular differentiation resulted in the specialization of chloroplasts (plastids) to serve specific metabolic functions within various tissue and cell types in addition to photosynthesis, for example, starch metabolism in roots and tubers (amyloplasts) and lipid metabolism during senescence and fruit ripening (chromoplasts)[76]. A simple plastid, the proplastid, evolved as the basic plastid form in undifferentiated meristematic tissues. The proplastids possess the ability to differentiate into any of the specialized forms during plant growth and development [77]. As with most plant cells, plastid differentiation is not terminal, and plants maintain the ability to interconvert the specialized organelles as required to match the metabolic demands of the cell. The coincidence of Toc159 appearance with the evolution of plastid differentiation has led to the proposal that the addition of the Toc159 receptors to the translocon conferred distinct preprotein selectivities on the individual translocon types, thereby giving rise to distinct protein targeting pathways [30, 32]. These pathways appear to be critical for the development of specific plastid types or specialized functions, perhaps by regulating the content and flux of protein trafficking into the organelles over the course of plant development.

*Arabidopsis* contains four isoforms of the Toc159 receptor: atToc159, atToc132, atToc120, and atToc90 [30, 78]. These isoforms appear to assemble with Toc75 and the two isoforms of Toc34, atToc33 and atToc34, to form structurally and functionally distinct Toc translocons (Figure 16.4) [9, 10].

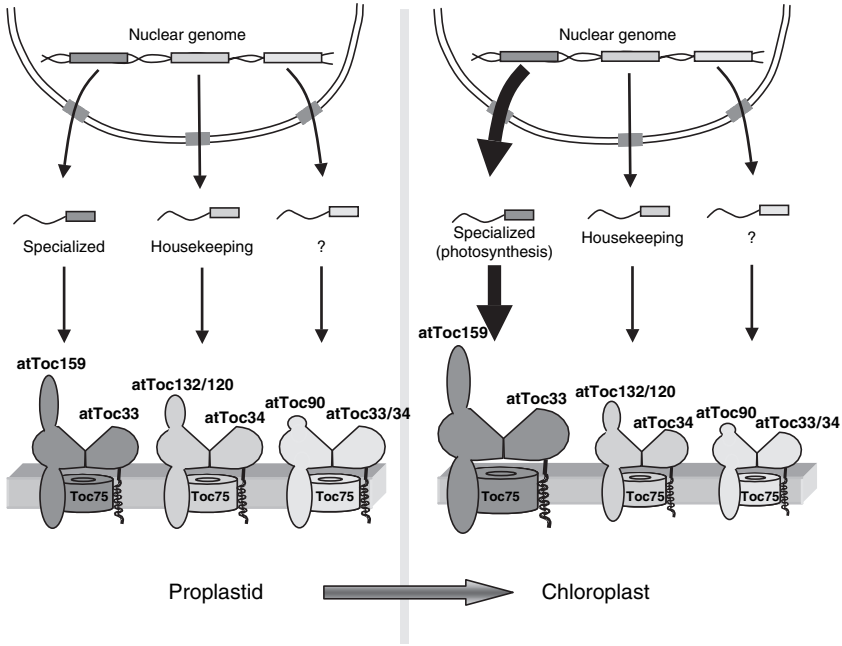



FIG. 16.4. The functional and structural diversity of Toc complexes. Small gene families encode the Toc GTPase receptors in *Arabidopsis*. The four members of the Toc159 family (atToc159, atToc132, atToc120, and atToc90) assemble with the two members of the Toc34 family (atToc33 and atToc34) and atToc75 to form structurally distinct Toc complexes. The atToc159 translocons appear to mediate the import of a subset of genes that are required for the development of chloroplasts, whereas the Toc132/Toc120 translocons are proposed to mediate the import of preproteins required for the housekeeping functions of all plastids. The structure and function of atToc90 complexes remains to be defined. The diversity of Toc complexes is essential to maintain the homeostasis of protein import during critical periods of plastid development.

AtToc159 is the most abundant receptor form in chloroplasts. Its expression is dramatically upregulated in green tissues, whereas atToc132 and atToc120 appear to be constitutively expressed at relatively lower levels in all tissues [9]. Null mutants of atToc159, *ppi2*, exhibit an albino phenotype and are unviable on soil [30]. However, *ppi2* plants are viable if supplemented with sucrose, indicating a primary deficiency in photosynthetic capacity within these plants. AtToc132 and atToc120 appear to be functionally redundant. However, double *attoc132/atoc120* mutants exhibit severe growth defects and are seedling lethal. This phenotype is not reversed by sucrose supplementation [9, 10]. These results indicate that atToc159 is primarily required for the development of photosynthetically competent chloroplasts, whereas

atToc132/atToc120 likely participate in the maintenance of plastid housekeeping functions, such as amino acid and lipid metabolism, and plastid replication and gene expression.

atToc132 and atToc159 are unable to rescue *ppi2* or *attoc132/attoc120* double mutant plants, respectively [9, 10]. Furthermore, these two receptor isoforms exhibit distinct preprotein selectivities in *in vitro* binding studies [9]. These observations, in conjunction with the fact that they form structurally distinct Toc translocons, indicate that they mediate different pathways for protein import into the organelle. The major structural variability among the Toc159 family members is in the length and sequence of their A-domains. This has led to the hypothesis that the A-domain dictates receptor specificity [30]. However, the A-domain does not appear to bind preproteins directly [40]. Therefore, its role in preprotein binding, if any, is likely to be through an indirect effect on the G-domain. AtToc33 and atToc34 also appear to contribute to the formation of distinct translocons (Figure 16.4). AtToc33 and atToc34 are enriched in atToc159 and atToc132 complexes, respectively [9]. Furthermore, they exhibit some preferences in preprotein binding consistent with their atToc159 and atToc132 counterparts [46]. However, *in vitro* studies demonstrate that the expression of atToc33 or atToc34 is capable of fully complementing null mutants in either gene [32]. Therefore, atToc33 and atToc34 exhibit significantly more functional overlap than that observed with the members of the Toc159 receptor family.

The existence of Toc complexes with distinct transit peptide specificities could explain why the definition of common targeting determinants within transit peptides has remained elusive. Transit peptides range in length from ~30–100 amino acids, and lack apparent similarities in their primary structures [79, 80]. Furthermore, biophysical analysis of synthetic transit peptides has failed to detect the formation of stable secondary structures in aqueous solution, indicating that transit peptides do not form structures analogous to the amphipathic $\alpha$-helices common to mitochondrial presequences [81, 82]. Although this apparent lack of common characteristics likely reflects the complexity of the targeting process itself, the common targeting determinants within transit peptides might only be detected within subsets of chloroplast preproteins that utilize a common import pathway.

The appearance of Toc34 and Toc75 as the original unit of the outer membrane import machinery provides a compelling argument that these two components define the minimal Toc translocon. This concept is further supported by import studies that demonstrate functional similarity in the ability of plastids to import preproteins from both simpler unicellular organisms and more complex higher plants [74]. The addition of the

Toc159 family to the translocon with the advent of plastid differentiation and the observation that these receptors define distinct translocons support the biochemical studies that Toc159 is primarily responsible for conveying preprotein specificity on Toc complexes [9, 40]. Although there are several genes encoding proteins with similarity to Toc75 in plants, only one of these appears to be involved in protein import [29]. Thus, the same Toc75 channel appears to be a common unit of all Toc translocons, consistent with the general concept of a nonselective protein-conducting channel that assembles with specific targeting receptors to generate a selective, unidirectional translocon.

## VII. Toc Complex Assembly

Nuclear genes encode all known translocon components. In addition to studies on the function of individual translocon components, recent investigations have begun to reveal the determinants and mechanism of Toc complex assembly. These investigations demonstrate that each of the three core Toc components is targeted to the outer membrane via a unique mechanism.

Toc34 is similar to a number of other outer membrane proteins in possessing an α-helical transmembrane domain that inserts into the lipid bilayer [15, 17]. As with other outer membrane proteins, Toc34 does not contain a transit peptide or other cleavable targeting signal. The information for targeting the receptor is contained within its transmembrane helix and in adjacent residues. Efficient insertion of Toc34 into the membrane relies on its intrinsic GTP binding; however, it is not clear whether nucleotide binding plays a direct role in insertion or an indirect role by stabilizing an insertion competent conformation. Studies with another model outer membrane protein, OEP14, suggest that Toc75 might provide the channel for insertion of this class of proteins into the membrane [11, 83].

Toc75 is unique among outer membrane proteins because it possesses a cleavable N-terminal transit peptide [84, 85]. This transit peptide is divided into two regions [84, 85]. The N-terminal region is functionally interchangeable with the transit peptides of proteins destined for the interior of the organelle [85]. Therefore, it appears that Toc complexes mediate at least some stages of Toc75 targeting to the chloroplast. The initial cleavage of the first segment of the transit peptide occurs in the stroma by the stromal processing peptidase [85], demonstrating that the N-terminal region of pre-Toc75 translocates into the stroma. This generates a size intermediate that remains associated with the envelope. The second region of the transit peptide contains a glycine-rich segment that appears to block complete

translocation of Toc75 across the envelope membranes and trigger its insertion into the outer membrane [85]. This segment of the transit peptide is cleaved later in the targeting process by an envelope-associated peptidase with similarity to the type 1 signal peptidases found in prokaryotes [86]. The mechanism of integration of the β-strands of Toc75 into the outer membrane is unknown. It also is not known whether Toc complexes play a direct role in Toc75 integration or simply provide the initial entry point for the polypeptide into the organelle. In mitochondria, the Tob/Sam proteins mediate β-barrel integration into the mitochondrial outer membrane [87]. Proteins similar to the Tob/Sam components have not been detected in genomic or proteomic analysis of chloroplast proteins, suggesting that this mechanism of β-barrel membrane integration is not conserved between the two organelles. As mentioned earlier, at least one protein with similarity to the Omp85 proteins of Gram-negative bacteria are present in chloroplasts [69], but it remains to be seen if they play direct roles in Toc75 biogenesis specifically or outer membrane biogenesis in general.

*In vitro* targeting studies with Toc159 demonstrate that its association with the outer membrane requires both Toc34 and Toc75 [88]. Binding of the receptor at the chloroplast surface appears to be mediated by an interaction between the G-domains of Toc34 and Toc159 [88]. These observations led to a third proposed role for Toc dimerization and GTPase activity; Toc159 targeting and assembly into Toc complexes (Figure 16.5). Studies using GTPase mutants and nonhydrolyzable GTP analogues demonstrate that the intrinsic GTPase activity of Toc159 is absolutely required

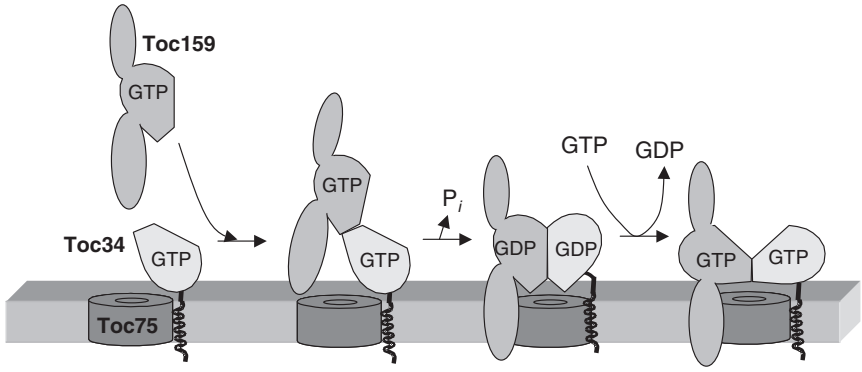



FIG. 16.5. The proposed role of Toc GTPase activity in the targeting and assembly of Toc159 into Toc complexes. Cytoplasmic Toc159 is targeted to Toc complexes by an interaction of G-domain with the G-domain of Toc34. GTP hydrolysis at both Toc159 and Toc34 is required for the insertion of the M-domain of Toc159 into the outer membrane and the assembly of the receptor into Toc complexes.

for insertion of the M-domain of the intact receptor into Toc complexes [48, 49]. In the absence of GTP hydrolysis, the receptor binds, but does not integrate into the outer membrane [48, 49]. Binding is mediated by G-domain dimerization between Toc159 and Toc34 [47, 48]. These observations suggest that dimerization mediates targeting of the receptor to the Toc complex, and subsequent GTP hydrolysis provides the energy for integration of the receptor into the translocon complex. As mentioned previously, atToc33 and atToc34 appear to be differentially enriched in Toc complexes containing atToc159 and atToc132 [9], respectively. Therefore, the formation of distinct Toc complexes could be dictated by preferential binding of the atToc159 receptors to atToc33 and atToc34. A population of Toc159 was detected in the cytoplasm concomitant with the discovery of the role of GTP in receptor targeting [50]. This led to the proposal that Toc159 might function as a cycling receptor whose docking, insertion, and release at the Toc translocon are regulated via a GTP binding and hydrolysis. To date, data supporting the existence of a cycling receptor and its relevance to preprotein targeting is lacking.

## VIII. Future Directions

Although the major players in Toc translocon function have been identified, the underlying mechanism of Toc function remains to be defined. The accumulated evidence supports the roles of Toc34 and Toc159 as an integrated preprotein receptor system at the Toc translocon [16, 19, 20, 39, 41]. Their GTPase activities appear to be critical in regulating Toc translocon assembly and function, providing additional novel roles for GTPases in regulating intracellular protein trafficking. The molecular details of GTPase function remain to be elucidated, and the availability of atomic resolution structures and the power of *Arabidopsis* genetics are providing experimental approaches that complement well-established *in vitro* import assays. The study of mutants containing alterations in GTPase *activity* and dimerization should reveal insights into the roles of these two properties of the receptors. In addition, the reverse genetic studies can address the relative contributions of the two receptors to preprotein recognition and the initiation of translocation at the Toc translocon. The molecular interactions between the receptors and the Toc75 channel also have to be defined. In particular, the mechanism by which preproteins are transferred from the surface receptors into the channel remains a mystery.

Another particularly exciting area of investigation centers on the roles of the distinct Toc complexes in plastid function and development. The substrate classes for each translocon remain to be discovered, and comparative studies among the translocons should provide important information on

the determinants of transit peptide-binding specificity. Investigations of the diversity of Toc complexes also must be expanded to include nonconventional translocons. Studies on the chloroplast proteome suggest the existence of hundreds of nucleus-encoded chloroplast proteins that lack detectable transit peptides [2]. The import of only a couple of these proteins has been studied, and it appears that they do not compete with stromal proteins for import. Therefore, it is likely that their targeting involves translocons that are distinct from those used by proteins containing standard transit peptides. It is of particular interest to determine if these pathways involve any of the known Toc translocons (e.g., Toc132/Toc120 translocons) or individual Toc components. In summary, it is clear that our concept of a single general import complex must be discarded and replaced with a picture of a diverse set of translocons that have evolved to reflect the complexity and diversity of plastid function within plants. Future studies not only will provide information on the mechanisms of novel import pathways but also will provide insight into the roles of these pathways in plastid development and responses to changes in cellular physiology during plant growth and development.

## References

1. Lopez-Juez, E., and Pyke, K.A. (2005). Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol* 49:557–577.
2. Jarvis, P. (2004). Organellar proteomics: chloroplasts in the spotlight. *Curr Biol* 14: R317–R319.
3. Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. *Plant Cell* 11:557–570.
4. Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci USA* 99:12246–12251.
5. Schnell, D.J., Blobel, G., Keegstra, K., Kessler, F., Ko, K., and Soll, J. (1997). A nomenclature for the protein import components of the chloroplast envelope. *Trends Cell Biol* 7:303–304.
6. Reumann, S., and Keegstra, K. (1999). The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. *Trends Plant Sci* 4:302–307.
7. Chua, N.H., and Schmidt, G.W. (1978). Post-translational transport into intact chloroplasts of a precursor to the small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc Natl Acad Sci USA* 75:6110–6114.
8. Highfield, P.E., and Ellis, R.J. (1978). Synthesis and transport of the small subunit of chloroplast ribulose-1,5-bisphosphate carboxylase. *Nature (Lond)* 271:420–424.
9. Ivanova, Y., Smith, M.D., Chen, K., and Schnell, D.J. (2004). Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. *Mol Biol Cell* 15:3379–3392.

10. Kubis, S., Patel, R., Combe, J., Bedard, J., Kovacheva, S., Lilley, K., Biehl, A., Leister, D., Rios, G., Koncz, C., and Jarvis, P. (2004). Functional specialization amongst the Arabidopsis Toc159 family of chloroplast protein import receptors. *Plant Cell* 16:2059–2077.
11. Tu, S.L., Chen, L.J., Smith, M.D., Su, Y.S., Schnell, D.J., and Li, H.M. (2004). Import pathways of chloroplast interior proteins and the outer-membrane protein OEP14 converge at Toc75. *Plant Cell* 16:2078–2088.
12. Schatz, G., and Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science* 271:1519–1526.
13. Schnell, D.J., and Hebert, D.N. (2003). Protein translocons: multifunctional mediators of protein translocation across membranes. *Cell* 112:491–505.
14. Hirsch, S., Muckel, E., Heemeyer, F., von Heijne, G., and Soll, J. (1994). A receptor component of the chloroplast protein translocation machinery. *Science* 266:1989–1992.
15. Kessler, F., Blobel, G., Patel, H.A., and Schnell, D.J. (1994). Identification of two GTP-binding proteins in the chloroplast protein import machinery. *Science* 266:1035–1039.
16. Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* 6:93–105.
17. Seedorf, M., Waegemann, K., and Soll, J. (1995). A constituent of the chloroplast import complex represents a new type of GTP-binding protein. *Plant J* 7:401–411.
18. Hinnah, S.C., Wagner, R., Sveshnikova, N., Harrer, R., and Soll, J. (2002). The chloroplast protein import channel toc75: pore properties and interaction with transit peptides. *Biophys J* 83:899–911.
19. Ma, Y., Kouranov, A., LaSala, S., and Schnell, D.J. (1996). Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. *J Cell Biol* 134:1–13.
20. Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. *J Cell Biol* 139:1677–1685.
21. Olsen, L.J., and Keegstra, K. (1992). The binding of precursor proteins to chloroplasts requires nucleoside triphosphates in the intermembrane space. *J Biol Chem* 267:433–439.
22. Theg, S.M., Bauerle, C., Olsen, L.J., Selman, B.R., and Keegstra, K. (1989). Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. *J Biol Chem* 264:6730–6736.
23. Young, M.E., Keegstra, K., and Froehlich, J.E. (1999). GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. *Plant Physiol* 121:237–244.
24. Akita, M., Nielsen, E., and Keegstra, K. (1997). Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. *J Cell Biol* 136:983–994.
25. Kessler, F., and Blobel, G. (1996). Interaction of the protein import and folding machineries in the chloroplast. *Proc Natl Acad Sci USA* 93:7684–7689.
26. Kourtz, L., and Ko, K. (1997). The early stage of chloroplast protein import involves Com70. *J Biol Chem* 272:2808–2813.
27. Nielsen, E., Akita, M., Davila-Aponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO J* 16:935–946.
28. Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. *Science* 266:1007–1012.

29. Baldwin, A., Wardle, A., Patel, R., Dudley, P., Park, S.K., Twell, D., Inoue, K., and Jarvis, P. (2005). A molecular-genetic study of the Arabidopsis Toc75 gene family. *Plant Physiol* 138:715–733.
30. Bauer, J., Chen, K., Hiltbunner, A., Wehrli, E., Eugster, M., Schnell, D., and Kessler, F. (2000). The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* 403:203–207.
31. Hust, B., and Gutensohn, M. (2006). Deletion of core components of the plastid protein import machinery causes differential arrest of embryo development in *Arabidopsis thaliana*. *Plant Biol (Stuttg)* 8:18–30.
32. Jarvis, P., Chen, L.-J., Li, H., Peto, C.A., Fankhauser, C., and Chory, J. (1998). An Arabidopsis mutant defective in the plastid general protein import apparatus. *Science* 282:100–103.
33. Hinnah, S.C., Hill, K., Wagner, R., Schlicher, T., and Soll, J. (1997). Reconstitution of a chloroplast protein import channel. *EMBO J* 16:7351–7360.
34. Marshall, J.E., DeRocher, A.E., Keegstra, K., and Vierling, E. (1990). Identification of heat shock protein hsp70 homologues in chloroplasts. *Proc Natl Acad Sci USA* 87:374–378.
35. Waegemann, K., and Soll, J. (1991). Characterization of the protein import apparatus in isolated outer envelopes of chloroplasts. *Plant J* 1:149–158.
36. Schleiff, E., Soll, J., Kuchler, M., Kuhlbrandt, W., and Harrer, R. (2003b). Characterization of the translocon of the outer envelope of chloroplasts. *J Cell Biol* 160:541–551.
37. Schleiff, E., Jelic, M., and Soll, J. (2003a). A GTP-driven motor moves proteins across the outer envelope of chloroplasts. *Proc Natl Acad Sci USA* 100:4604–4609.
38. Constan, D., Patel, R., Keegstra, K., and Jarvis, P. (2004). An outer envelope membrane component of the plastid protein import apparatus plays an essential role in Arabidopsis. *Plant J* 38:93–106.
39. Schleiff, E., Soll, J., Sveshnikova, N., Tien, R., Wright, S., Dabney-Smith, C., Subramanian, C., and Bruce, B.D. (2002). Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, toc34. *Biochemistry* 41:1934–1946.
40. Smith, M.D., Rounds, C.M., Wang, F., Chen, K., Afitlhile, M., and Schnell, D.J. (2004). atToc159 is a selective transit peptide receptor for the import of nucleus-encoded chloroplast proteins. *J Cell Biol* 165:323–334.
41. Sveshnikova, N., Soll, J., and Schleiff, E. (2000b). Toc34 is a preprotein receptor regulated by GTP and phosphorylation. *Proc Natl Acad Sci USA* 97:4973–4978.
42. Bolter, B., Soll, J., Schulz, A., Hinnah, S., and Wagner, R. (1998). Origin of a chloroplast protein importer. *Proc Natl Acad Sci USA* 95:15831–15836.
43. Bolter, B., May, T., and Soll, J. (1998). A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. *FEBS Lett* 441:59–62.
44. Chen, K., Chen, X., and Schnell, D.J. (2000). Initial binding of preproteins involving the Toc159 receptor can be bypassed during protein import into chloroplasts. *Plant Physiol* 122:813–822.
45. Sun, Y.J., Forouhar, F., Li Hm, H.M., Tu, S.L., Yeh, Y.H., Kao, S., Shr, H.L., Chou, C.C., Chen, C., and Hsiao, C.D. (2002). Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nat Struct Biol* 9:95–100.
46. Jelic, M., Soll, J., and Schleiff, E. (2003). Two Toc34 homologues with different properties. *Biochemistry* 42:5906–5916.
47. Weibel, P., Hiltbrunner, A., Brand, L., and Kessler, F. (2003). Dimerization of Toc-GTPases at the chloroplast protein import machinery. *J Biol Chem* 278:37321–37329.

48. Smith, M.D., Hiltbrunner, A., Kessler, F., and Schnell, D.J. (2002). The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *J Cell Biol* 159:833–843.
49. Bauer, J., Hiltbrunner, A., Weibel, P., Vidi, P.A., Alvarez-Huerta, M., Smith, M.D., Schnell, D.J., and Kessler, F. (2002). Essential role of the G-domain in targeting of the protein import receptor atToc159 to the chloroplast outer membrane. *J Cell Biol* 159:845–854.
50. Hiltbrunner, A., Bauer, J., Vidi, P.A., Infanger, S., Weibel, P., Hohwy, M., and Kessler, F. (2001). Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. *J Cell Biol* 154:309–316.
51. Kessler, F., and Schnell, D.J. (2004). Chloroplast protein import: solve the GTPase riddle for entry. *Trends Cell Biol* 14:334–338.
52. Jelic, M., Sveshnikova, N., Motzkus, M., Horth, P., Soll, J., and Schleiff, E. (2002). The chloroplast import receptor Toc34 functions as preprotein-regulated GTPase. *Biol Chem* 383:1875–1883.
53. Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. *Annu Rev Biochem* 70:755–775.
54. Lee, K.H., Kim, S.J., Lee, Y.J., Jin, J.B., and Hwang, I. (2003). The M domain of atToc159 plays an essential role in the import of proteins into chloroplasts and chloroplast biogenesis. *J Biol Chem* 278:36794–36805.
55. Becker, T., Jelic, M., Vojta, A., Radunz, A., Soll, J., and Schleiff, E. (2004b). Preprotein recognition by the Toc complex. *EMBO J* 23:520–530.
56. Sveshnikova, N., Grimm, R., Soll, J., and Schleiff, E. (2000a). Topology studies of the chloroplast protein import channel Toc75. *Biol Chem* 381:687–693.
57. Kouranov, A., Chen, X., Fuks, B., and Schnell, D.J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J Cell Biol* 143:991–1002.
58. Clark, S.A., and Theg, S.M. (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol Cell Biol* 8:923–934.
59. Schnell, D.J., and Blobel, G. (1993). Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. *J Cell Biol* 120:103–115.
60. Wu, C., Seibert, F.S., and Ko, K. (1994). Identification of chloroplast envelope proteins in close proximity to a partially translocated chimeric precursor protein. *J Biol Chem* 269:32264–32271.
61. Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., and Schleiff, E. (2004a). Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. *Mol Biol Cell* 15:5130–5144.
62. Ko, K., Bornemisza, O., Kourtz, L., Ko, Z.W., Plaxton, W.C., and Cashmore, A.R. (1992). Isolation and characterization of a cDNA clone encoding a cognate 70-kDa heat shock protein of the chloroplast envelope. *J Biol Chem* 267:2986–2993.
63. May, T., and Soll, J. (2000). 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* 12:53–64.
64. Qbadou, S., Becker, T., Mirus, O., Tews, I., Soll, J., and Schleiff, E. (2006). The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. *EMBO J* 25:1836–1847.
65. Nakrieko, K.A., Mould, R.M., and Smith, A.G. (2004). Fidelity of targeting to chloroplasts is not affected by removal of the phosphorylation site from the transit peptide. *Eur J Biochem* 271:509–516.

66. Sohrt, K., and Soll, J. (2000). Toc64, a new component of the protein translocon of chloroplasts. *J Cell Biol* 148:1213–1221.
67. Rosenbaum Hofmann, N., and Theg, S.M. (2005). Toc64 is not required for import of proteins into chloroplasts in the moss Physcomitrella patens. *Plant J* 43:675–687.
68. Delwiche, C.F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* 154:S164–S177.
69. Reumann, S., Inoue, K., and Keegstra, K. (2005). Evolution of the general protein import pathway of plastids (review). *Mol Membr Biol* 22:73–86.
70. Reumann, S., Davila-Aponte, J., and Keegstra, K. (1999). The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. *Proc Natl Acad Sci USA* 96:784–789.
71. Voulhoux, R., Bos, M.P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299:262–265.
72. Hofmann, N.R., and Theg, S.M. (2003). Physcomitrella patens as a model for the study of chloroplast protein transport: conserved machineries between vascular and non-vascular plants. *Plant Mol Biol* 53:621–632.
73. McFadden, G.I., and van Dooren, G.G. (2004). Evolution: red algal genome affirms a common origin of all plastids. *Curr Biol* 14:R514–R516.
74. Steiner, J.M., Yusa, F., Pompe, J.A., and Loffelhardt, W. (2005). Homologous protein import machineries in chloroplasts and cyanelles. *Plant J* 44:646–652.
75. Jakowitsch, J., Neumann-Spallart, C., Ma, Y., Steiner, J., Schenk, H.E., Bohnert, H.J., and Loffelhardt, W. (1996). *In vitro* import of pre-ferredoxin-$NADP^+$-oxidoreductase from Cyanophora paradoxa into cyanelles and into pea chloroplasts. *FEBS Lett* 381:153–155.
76. Bedard, J., and Jarvis, P. (2005). Recognition and envelope translocation of chloroplast preproteins. *J Exp Bot* 56:2287–2320.
77. Vothknecht, U.C., and Westhoff, P. (2001). Biogenesis and origin of thylakoid membranes. *Biochim Biophys Acta* 1541:91–101.
78. Hiltbrunner, A., Grünig, K., Alvarez-Huerta, M., Infanger, S., Bauer, J., and Kessler, F. (2004). AtToc90, a new GTP-binding component of the Arabidopsis chloroplast protein import machinery. *Plant Mol Biol* 54(3):427–440.
79. Bruce, B.D. (2000). Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol* 10:440–447.
80. Zhang, X.P., and Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci* 7:14–21.
81. Bruce, B.D. (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochim Biophys Acta* 1541:2–21.
82. von Heijne, G., and Nishikawa, K. (1991). Chloroplast transit peptides: the perfect random coil? *FEBS Lett* 278:1–3.
83. Tu, S.L., and Li, H.M. (2000). Insertion of OEP14 into the outer envelope membrane is mediated by proteinaceous components of chloroplasts. *Plant Cell* 12:1951–1960.
84. Tranel, P.J., Froehlich, J., Goyal, A., and Keegstra, K. (1995). A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. *EMBO J* 14:2436–2446.
85. Tranel, P.J., and Keegstra, K. (1996). A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. *Plant Cell* 8:2093–2104.
86. Inoue, K., Baldwin, A.J., Shipman, R.L., Matsui, K., Theg, S.M., and Ohme-Takagi, M. (2005). Complete maturation of the plastid protein translocation channel requires a type I signal peptidase. *J Cell Biol* 171:425–430.

87. Paschen, S.A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* 426:862–866.
88. Wallas, T.R., Smith, M.D., Sanchez-Nieto, S., and Schnell, D.J. (2003). The roles of toc34 and toc75 in targeting the toc159 preprotein receptor to chloroplasts. *J Biol Chem* 278:44289–44297.

# 17

# *The Role of the Tic Machinery in Chloroplast Protein Import*

PHILIPP BENZ • JÜRGEN SOLL • BETTINA BÖLTER

*Department Biologie I*
*Ludwig-Maximilians-Universität München*
*Menzinger Str. 67, Munich D-80638, Germany*

## I. Abstract

Plastids are a heterogeneous family of organelles found ubiquitously in plant cells. Most prominent are the chloroplasts which carry out essential functions such as photosynthesis, the biosynthesis of fatty acids, as well as amino acids.

Like mitochondria, chloroplasts are derived from an endosymbiotic event. They are believed to have evolved from an ancient cyanobacterium which has been engulfed by an early eukaryotic mitochondriate ancestor. During evolution, the plastid genome has been greatly reduced, and most of the genes have been transferred to the host nucleus. Consequently, most plastid proteins are translated in the cytosol and must be posttranslationally imported into the organelle. Thus, chloroplast development is dependent on the coordinated expression of both plastid and nuclear-encoded genes. Especially import into the chloroplast, which is mediated by complex molecular machines in the outer (Toc complex) and inner (Tic complex) envelope of chloroplasts, respectively, and assembly into functional complexes have to be tightly regulated.

This chapter deals with the translocon of the inner envelope membrane (IEM), the Tic complex. Seven components have been annotated so far to exert a function in protein translocation across the inner envelope.



DOI: 10.1016/S1874-6047(07)25017-6

Here, we summarize what is known today about the single Tic constituents, take a glimpse at evolutionary aspects and discuss a possible regulation mechanism of Tic.

## II. Introduction

Chloroplasts are the prominent organelles of green plant tissue. According to the widely accepted endosymbiotic theory [1] they originated from an ancient cyanobacterium, which was engulfed by a eukaryotic host cell. The metamorphosis of the autonomous cyanobacterium into a cell organelle included the transfer of the cyanobacterial genes to the host nucleus. About 98% of all chloroplast genes are now found in the nuclear genome. Consequently, the vast majority of chloroplast proteins is synthesized on cytosolic ribosomes and has to be transported posttranslationally into the plastid. This includes specific targeting to the organelle, binding to receptor proteins on the organellar surface and translocation across the double envelope membranes. Most of the proteins destined for a plastidic subcompartment are synthesized as precursor proteins with a cleavable N-terminal presequence. Targeting sequences reveal only little similarity at the level of primary sequence or length [2], but all contain predominantly positively charged and hydroxylated amino acids such as threonine and serine [3]. Proteins which are located in the outer envelope use a different, yet mostly unknown, mechanism for targeting and insertion [4]. Translocation itself is mediated by two complex machineries, called *T*ranslocon at the *o*uter envelope of *c*hloroplasts (Toc) and *T*ranslocon at the *i*nner envelope of *c*hloroplasts (Tic).

The characteristics of Toc are extensively described by D. Schnell in this volume. Here, we will delineate the knowledge about Tic, which is still rather scarce at this point. In general, single subunits have been characterized by biochemical approaches, especially focusing on interaction with precursor proteins. For each member convincing evidence has been provided of it being a Tic component, even though their specific function in most cases remains to be clarified. The same applies to the exact composition of the complex under different conditions and the regulation of the import process across the inner envelope. Due to the growing number of available mutant lines of *Arabidopsis*, also genetic approaches shed further light on the possible function of Tic constituents. The analysis of knockout or antisense mutants of the respective Tic genes is a very valuable tool for gaining more insight into protein translocation into chloroplasts. We will focus on the characterization of the single subunits and try to fit all known facts into a larger picture of Tic function and regulation.

## III. Tic110: The Translocation Channel

Tic110 was the first component of the inner envelope translocon to be identified and still is the best-studied subunit today.

Its most prominent feature is its abundance in the inner envelope membrane (IEM), which made it an early target for chloroplast research and of course a good marker protein for this membrane fraction in chloroplast preparations [5–7]. Additionally, analysis of the Tic110 expression pattern showed that it is present throughout development in various green and nongreen tissue types, pointing out its importance in the general import pathway [8, 9].

Cloning as well as molecular and biochemical characterization of the protein from pea (*Pisum sativum*) confirmed that Tic110 is an integral inner membrane protein containing a 37-amino acid cleavable presequence for chloroplast localization, and suggested the presence of two hydrophobic transmembrane (TM) α-helices in its extreme N-terminus followed by a long (>97 kDa) hydrophilic domain [10, 11]. However, the topology of this C-terminal domain as well as Tic110's functional properties have been a matter of debate (see below). DNA sequence analysis and database queries showed that the Tic110 gene is well conserved among higher plants and shares no obvious homologies to any other protein family (except to itself in certain domains pointing to partial gene duplication events during evolution) [12].

Tic110 was initially characterized as a Tic component when it was found associated with a late translocation intermediate of precursor protein [13, 14]. These late translocation intermediates were observed when a nuclear-encoded preprotein (usually preSSU) destined for the chloroplast was imported *in vitro* under low ATP conditions ($\leq$100 μM), which slowed the translocation process of this precursor down on its way through the envelope membranes. Subsequent cross-linking and purification of the resulting complex led to copurification of Tic110 with the preprotein.

The high abundance and the size of Tic110 suggests a role as a central subunit of the IEM translocon with complex regulation and functional properties [15]. Several studies have found this assumption to be true, associating Tic110 with at least four different features necessary for effective preprotein import:

1. Constituting the translocon pore [16]
2. Reacting to the redox state of the chloroplast (see sections VII–IX) [17–19]
3. The recruitment of chaperones to the import site (whether directly or indirectly via Tic40) [10,20–22]
4. The formation of Toc–Tic joint translocation sites [23]

It is generally accepted that preprotein import occurs simultaneously through the OEM and IEM of the chloroplast in distinct EM-visible patches where the two envelope membranes come in physical proximity, called contact sites [13, 24]. Toc and Tic complexes are believed to interact in these areas to enable a more direct translocation of preproteins from the cytosol to the chloroplast stroma. However, the exact subunit composition is still elusive.

In addition, it is not clear whether these sites are permanently present or associate only on precursor translocation. Despite these open questions, the theory has been confirmed in many experiments in which Tic110 was shown to coimmunoprecipitate with Toc components (and vice versa) from cross-linked complexes in the presence of precursors [11, 20]. It has to be noted though that only a minor portion of the Tic110 pool (as well as Toc) is assembled into these high molecular weight complexes at any given time [9, 21] and that the connection from Tic110 to Toc is most likely not direct but mediated by precursor proteins and/or other intermembrane space (IMS)-localized factors like Tic22 (see below) and Toc12, but clearly Tic110 is required for a functional import machinery composed of both translocons.

The exact conformations of Tic110 as well as its topology in the IEM still need to be established, but from the experimental data presented earlier, it can be concluded that parts of Tic110 are exposed to the IMS. The short loop of only a few amino acids connecting the two predicted N-terminal transmembrane (TM)-helices alone would probably be insufficient to fulfil this role, therefore large portions of the long C-terminus should be exposed to the IMS, as proposed in the topology model by Lübeck *et al.* (1996) [11]. This conclusion was based on trypsin treatment of inner envelope vesicles as well as limited proteolysis of intact chloroplasts using an experimental setup which did not allow the protease to reach the stroma. Moreover, Heins *et al.* (2002) showed that the C-terminus of Tic110 alone could integrate into artificial liposomes [16] (see below). Jackson *et al.* (1998) challenged this view, arguing that Tic110 was not degraded by added protease under their experimental conditions [25]. Additionally, Inaba *et al.* (2003) found that an N-terminally truncated atTic110 construct expressed in *Arabidopsis* was not inserted into the IEM [26]. These results led the authors to conclude that the C-terminus rather forms a hydrophilic stromal domain which may coordinate protein import at this side of the Tic complex.

Another important step occurring at the IEM is the direct binding to incoming transit peptides. Using hexahistidine-tagged preprotein fusions (or the respective mature forms) immobilized on a $Ni^{2+}$-NTA resin together with *in vitro* translated [$^{35}$S] atTic110 in a binding assay, Inaba *et al.* (2003) demonstrated that Tic110 indeed has the ability to bind the

transit peptide of preproteins directly [26]. Exploiting a series of N-terminal deletion constructs of Tic110 as competitors in the binding assay, the group could even narrow down the transit peptide-binding domain roughly to a region between amino acids 185 and 370. They hypothesized that this domain was located adjacent to the translocon pore and therefore in a promising position to bind the TP directly after transit through the Tic channel. Other stromal proteins could very likely then ''take over'' the precursor and maybe use the Tic110 C-terminus as a platform for coordinated action [26]. In their subsequent work, Inaba *et al.* (2005) also claimed that amino acids 93–602 of atTic110 are important for the assembly into Toc–Tic complexes [9]. Even though it was still argued that this region was part of the stromal domain, both results would also be in line with a model in which parts of the C-terminus face the IMS.

Early chloroplast import studies already demonstrated that cytoplasmically (or *in vitro*) synthesized polypeptides are imported into the chloroplast in an ATP-dependent process [27] in which the energy is not used to generate a membrane potential for driving the import reaction but that exerts its effect on a stromal ATPase with a different function [28]. Subsequently, at least two proteins (Cpn60 and ClpC) have been reported to associate specifically with the Tic complex that could represent those factors postulated a decade earlier [10, 22].

*C*ha*p*eroni*n* of 60 kDa (Cpn60) represents a stromal oligomeric ATPase homologous to bacterial GroEL (Hsp60 family), which acts in the folding of proteins in its central cavity [29, 30]. It was initially implicated with preprotein translocation when it was found to be the major protein coimmunoprecipitating with Tic110 in Triton X-100 solubilized chloroplasts (without added precursors [10]). The interaction was further shown to be specific and ATP-sensitive (detectable only in the absence of ATP). In the same ATP-dependent manner, both Tic110 and Cpn60 associated with the processed (mature) form of a newly imported (urea-denatured) protein (mSSU). Binding of the precursor form (preSSU) was also detected, but this interaction was much less ATP-dependent and by that in all probability a direct interaction with Tic110, whereas the interaction of Tic110 with mSSU was mediated by either Cpn60 (as claimed by the authors) or the stromal processing peptidase, which obviously has to be present at this stage as well [31].

ClpC (or Hsp93) is a second stromal chaperone shown to be involved in the late translocation events [20, 22]. It belongs to the Hsp100 class of chaperones and represents a functional homologue to prokaryotic ClpA, being the regulatory ATPase subunit of the serine protease Clp (*c*aseino*l*ytic *p*rotease [32]). ClpA/C and their proteolytic counterpart ClpP associate to form oligomeric ring-like complexes which have been implicated

in intracellular degradation and substrate turnover [33]. Interestingly, the ClpC protein is also able to function as an independent molecular chaperone—promoting protein unfolding and disassembly of protein aggregates and complexes [34]. That ClpC also supports the translocation process at the Tic complex was put into focus by Nielsen *et al.* (1997), who showed that components of both Tic (meaning: Tic110) and Toc could be found in a complex with ClpC under import conditions (and even when precursor was not present) [22]. Using a coimmunoprecipitation approach with chloroplasts solubilized in a mild detergent, the interaction was again demonstrated to be specific and destabilized by ATP, analogous to the results obtained for Cpn60. An important difference became obvious when the interaction with precursor proteins was analyzed in a time course import experiment: while Cpn60 was found to predominantly associate directly with the mature form of translocated proteins, ClpC interaction could only be detected with the precursor, implicating that the binding process takes place before the stromal processing peptidase and Cpn60 come into action and very shortly after the preprotein emerges from Tic110. Analysis of T-DNA knockout mutants of the predominantly expressed ClpC isoform in *Arabidopsis* (*Hsp93-V*) confirmed its function in protein import [35]. Homozygous mutant plants showed a pale phenotype, defects in chloroplast biogenesis, and a significant reduction in import rate *in vitro*.

Surprisingly, no specific interaction of any chloroplastidic Hsp70 homologue with the import machinery could be detected [20, 22]. A recruitment of a stromal Hsp70 by Tic110—analogous to the Tim44–mtHsp70 interaction in the mitochondrial PAM motor complex [36]—could have been envisaged but does not seem to exist. Instead, published data argue in favor of ClpC being (at least part of) the ''import motor,'' probably fixing the preprotein and preventing any retrograde movement in a Brownian ratchet manner [37], and maybe in cooperation with Tic110.

The most important function assigned to Tic110 is its role as the Tic channel [16]. Being the pore component of Tic fits nicely to the large amount of available data which put Tic110 in a very central position of preprotein translocation across the IEM (see above). Direct interactions with transit peptides of incoming precursors, the Toc translocon, most of the other Tic subunits as well as the stromal chaperones (all reviewed in [15]) would be very well in line with this proposed function.

Heins *et al.* (2002) demonstrated that heterologously expressed Tic110 from pea (as well as a construct without the N-terminal TM-helices) reconstituted into liposomes formed an ion channel [16]. It had a high conductance and was shown to prefer cations over anions. This seems appropriate for the usually positively charged transit peptides of translocated precursors and fits very well to the unusually high abundance of

negatively charged amino acids in the Tic110 protein [16]. The Tic110 channel displayed a specific block after the addition of purified transit peptide at the *trans* side of the lipid bilayer. The electrophysiological properties indicated a pore size of ~15–20 Å, which is similar to the pore size of the OE channel Toc75 and large enough to allow the passage of an unfolded polypeptide chain. Furthermore, Heins *et al.* identified a channel with very similar properties (high conductance, cation-selectivity, and voltage-dependence) after fusion of inner envelope vesicles with a planar bilayer, supporting the suggestion that both channels are identical.

Again, the debate over Tic110 being a channel protein or rather a scaffold for the coordination of purely stromal events brought back the question about its topology. As mentioned earlier, Inaba *et al.* (2003) found their soluble constructs of the atTic110 C-terminus protease-indigestible [26]. On the other hand, Heins *et al.* (2002) clearly demonstrated that the C-terminus of Tic110 is able to form a channel independent of the N-terminal TM-helices [16]. Moreover, after reconstitution into liposomes, Tic110 could be digested with protease resulting in a similar pattern as Lübeck *et al.* (1996) had shown for the native protein in intact chloroplasts [11]. These results seem mutually exclusive but may indicate that the topology of Tic110 is actually more complex than previously assumed (probably depending on the environmental or experimental conditions). Clearly, further structural investigations are needed to finally solve this issue.

More recently, the availability and detailed analysis of T-DNA knockout mutants in *Arabidopsis thaliana* has provided valuable insights into the Tic subunits and their respective functions *in vivo*. Concerning Tic110, Inaba *et al.* (2005) and Kovacheva *et al.* (2005) were able to confirm and extend our knowledge about chaperone recruitment and Tic complex assembly [9, 35]. Most importantly, T-DNA knockouts of the Tic110 gene were shown to have an embryo-lethal phenotype in the homozygous state. Mutant embryos did not progress beyond the globular stage and had a ''raspberry''-like phenotype (compare [38]). Interestingly, a very similar phenotype had been described for one of Tic110's interaction partners (Cpn60α [39]) as well as for a gene of unknown function predicted also to be localized in the chloroplast, called Raspberry 3 (Rsy3 [40]). These findings highlight that Tic-mediated action is essential from the very beginning of chloroplast development and, more general, that plastid biogenesis and embryo development seem to be tightly linked.

Equally interesting is the fact that even the heterozygous knockout plants already showed a phenotype, indicating that mutations in the *Tic110* gene are semidominant [35]. This effect is rather rare and underlines the importance of Tic110. From the mentioned high abundance of Tic110 in the IEM alone, one could have anticipated that a drop even by ~50% of functional

transcripts has no effect on plants, but the opposite seems to be true. Heterozygous plants were shown to have a subtle but obvious chlorotic phenotype and the import rates in chloroplasts isolated from these plants were significantly reduced [9, 35]. Deleterious effects were also observed when full length Tic110 or N- and C-terminally truncated constructs, respectively, were overexpressed in *Arabidopsis*. The most plausible explanation for these results is the presence of a highly dynamic but sensitive stoichiometry of subunits within the complex. Obviously, the amount of Tic110 integrated into the translocon is directly dependent on the overall amount of expressed protein. Too little or too much immediately disrupts the stoichiometric order of the system or, in the case of the truncated constructs, displaces functional subunits from the complexes [9].

Mutant plants of two different genes, which are both viable, can be crossed to study the effect of double mutations in one genetic background. This can be used to assess functional relationships between the analyzed proteins. If the new phenotype is a sum of severity of both single phenotypes, the proteins are generally believed to function in different pathways. However, if no phenotypic additivity becomes obvious, but one phenotype masks the other or the double mutants rather show signs of suppression, there is reason to believe that they are involved in a common process. This effect is called ''epistasis.'' Kovacheva *et al.* (2005) demonstrated that epistasis comes into effect when double mutants of *attic110* and *athsp93-V* (see above) as well as another Tic component, called atTic40, are generated [35]. This indication for a close functional cooperation has only been the latest in a long row of experiments showing that Tic40 indeed plays a prominent role in preprotein translocation across the IEM.

## IV. Tic40: The Cochaperone

Tic40 had initially been identified as two envelope polypeptides of ~44 and ~42 kDa, respectively, involved in chloroplast protein import by virtue of their physical proximity to translocating proteins [14]. Conflicting results did not allow the clear assignment of the protein(s) to either the inner or the outer envelope membrane, reflected by the given name *c*hloroplast *i*nner *m*embrane/*c*hloroplast *o*uter *m*embrane protein of 44 kDa (Cim/Com44). After subsequent identification of a first seemingly full-length cDNA clone from *Brassica napus*, the protein was renamed Toc36 due to the finding that the *in vitro* translated peptide was only attached to the chloroplast but neither processed nor imported [41]. Finally, Stahl *et al.* (1999) settled this issue when they identified a pea cDNA clone with an extended 5′-terminus

(compared to *bnToc36)* and demonstrated by N-terminal sequencing of the native protein isolated from detergent-solubilized envelope membranes that the clone originally identified by Ko *et al.* (1995) did indeed not represent a full-length clone but lacked the N-terminal signal peptide, which had probably been the main reason for its mistargeting to the OE [42]. Additionally, the N-terminal sequencing revealed that the 42 kDa polypeptide originally thought to be a second family member actually represents only a proteolytic fragment of Tic40.

Secondary structure prediction algorithms indicated that Tic40 contains a single TM span within the extreme N-terminal region of the protein, anchoring it to the envelope membrane. Stahl *et al.* (1999) also showed that the C-terminus of Tic40 is protease resistant and thereby most likely oriented towards the stroma. Using biochemical fractionation of purified chloroplasts as well as immunogold labeling of ultrathin sections from pea leaves, the exclusive localization of Tic40 to the IEM could be demonstrated.

As for the possible function of Tic40, sequence alignments and BLAST searches indicated that the best conserved regions are around the TM domain and in the C-terminal half of the hydrophilic domain. Especially the region from residues 380 to 438 of psTic40 shows a high homology (~40% identity) to the C-terminal globular domain shared by the cochaperones *H*sp70-*i*nteracting *p*rotein (Hip) and *H*sp70 and Hsp90-*o*rganizing *p*rotein (Hop) [42–44].

Hip represents one of several cochaperones that are known to regulate the activities of the Hsp70 chaperone family [45, 46]. It appears to stabilize Hsp70 in its ADP-bound conformation, prolonging its association with the substrate, but was also shown to possess chaperone activity on its own, independent of Hsp70. Hop then seems to subsequently mediate the cooperation and physical association between the Hsp70/Hsp40/Hip complex and Hsp90 (necessary to generate the so-called ''intermediate complex'' in the Hsp90 chaperone cycle [47, 48]), but a Hop protein from yeast, Sti1, has also been shown to bind to Hsp104, another Hsp100 family member [49]. Also involved in the binding processes of Hip and Hop to Hsp70, Hsp90, and Hsp104, respectively, is a structural motif called *t*etratrico-*p*eptide *r*epeat (TPR; for review see [50]). These domains are degenerate 34-amino acid repeats often arranged in tandem arrays, each forming two antiparallel α-helices, and they are used ubiquitously to mediate a variety of protein–protein interactions (generally with non-TPR proteins). In addition to Hip and Hop, TPRs are present in a number of functionally unrelated proteins.

For Tic40, further tertiary structure predictions indicated that immediately preceding the Hip/Hop domain, seven α-helices are found that form a

structure with similarity to TPR domains. Experimental evidence supporting this prediction was gained from the finding that an antibody raised against TPR1, a protein containing a single TPR domain, could cross-react with Tic40 [43].

Taken together, the structural predictions strongly suggest that Tic40 may have the function of an IEM-anchored cochaperone. Theoretically, Tic40 could either associate with the unfolded preproteins as they emerge from the Tic channel (analogous to or in coordination with Cpn60) or modulate the activity of a close chaperone, potentially Hsp93. Experimental data supporting the former notion is rather scarce. In the previously mentioned experiments performed by Kessler and Blobel (1996) which identified Cpn60 as a major component coimmunoprecipitating with Tic110, another minor signal of about 40 kDa had been detected, which was argued to possibly represent IAP36 (now Tic40), but no further data were presented to clarify this point [10]. Additionally, time course import experiments presented by Chou *et al.* (2003) showed predominant (but not exclusive) interaction with the mature form of SSU [43]. All these data could link Tic40 to the posttranslocational chaperone-mediated folding processes, but other, more recent, indications gained from biochemical as well as genetic data put Tic40 rather in an earlier position as part of the Tic motor complex, together with Tic110 and Hsp93/ClpC.

Wu *et al.* (1994) had already shown Tic40 to be in a complex associated with late import intermediates, a stage of import halted before processing and folding occurs [14]. Furthermore, via chemical cross-linking of whole chloroplasts with DSP or in isolated inner envelope vesicles with DSP and even the 0-Å cross-linker $CuCl_2$ [42], a very close connection of Tic40 to Tic110 has been reliably established. In the same cross-linking studies, Chou *et al.* (2003) also investigated the presence of Toc subunits or stromal chaperones implicated in protein translocation in the Tic40-immunoprecipitates. These experiments revealed an association with Toc75 and Hsp93/ClpC but not with Toc159, Hsp70, or Cpn60, speaking in favor of the latter hypothesis. Moreover, when Chou *et al.* (2003) characterized the import efficiency of chloroplasts from mutant *tic40* plants, binding of preproteins to the chloroplasts seemed normal, but striking differences became obvious when import was performed under high ATP conditions (5 mM). Total import of the analyzed precursors was reduced by 25–40%, but, even more interesting, significant quantities of precursor as well as small amounts of mature forms were detectable in the posttranslational supernatant even though chloroplasts had been reisolated after the precursor-binding step. Additionally, immunoblots from whole chloroplasts showed that unprocessed forms of chloroplastidic proteins could be found also *in vivo*. The authors concluded from these results that the absence of

Tic40 led to defects in the ability to interact with preproteins during late import steps and in translocation to the stroma, thereby releasing them back into the cytosol or import medium, respectively [43]. This proposal would draw a nice parallel to the already discussed ''motor'' function of ClpC (see above), and could be done in a concerted action.

In consequence of the available data, a close connection of Tic40, Tic110, and the chaperone ClpC can be anticipated. The reported similarity in chloroplast phenotypes from the respective mutants, concerning, for example, the defects in import efficiency or chloroplast ultra structure [35, 43], as well as the accumulated biochemical data, place all three components at the same stage in preprotein import. The most intriguing possibility for a concerted action, of course, would be the role as the Tic import motor. The fact that the *tic40* knockout plants are still viable though, while *tic110* mutants are embryo-lethal, shows that Tic40 function does not seem to be essential for plant development in general. If this was the case, Tic40 could rather fulfill an accessory function, maybe in modulation or optimization of the ATPase activity of Hsp93.

However, another interesting side effect of the *tic40* phenotype was a significant reduction in the number of chloroplasts in the mutant plants [35, 51]. It was further demonstrated that undeveloped plastids were present, which were then degraded more extensively in the vacuoles as compared to wild-type. These results could give an indication of the fact that Tic40 may also be an important factor for chloroplast biogenesis and survival in the cell [51].

## V. Tic20: A Putative Channel Protein

At first sight, Tic110 and another IEM protein, called Tic20, could not be more different, but still both proteins have been implicated with the role of being a translocation channel [16, 52]. However, the idea of Tic20 being able to form a channel has not yet been proven experimentally and is based mainly on sequence comparison studies and structural predictions [53].

The initial detection of Tic20 has to be contributed to Ma *et al*. (1996), who detected the protein cross-linked to a modified version of preSSU in an intermediate stage of import [54]. After subsequent identification of a full-length cDNA clone from pea, BLAST searches revealed a distant sequence similarity to two prokaryotic branched-chain amino acid transporters (AZLC of *Bacillus subtilis* and LivH of *Methanococcus jannaschii* [53]). A similar sequence similarity had also been detected for the mitochondrial channel proteins Tim23, Tim22, and Tim17 [55]. A structural feature common to Tic20 and the Tim proteins is, for example, the prediction

of four TM-domains, embedding them deeply into the inner membrane of mitochondria and chloroplasts, respectively. Using a covalent label-transfer cross-linking study in order to map the sequence of events during import, Kouranov and Schnell (1997) found Tic20 (and Tic22) in intermediate to late stages, with Tic22-binding preproteins before Tic20 [56]. Interestingly, Tic110 could not be detected in this study. The presence of a close connection of Tic20 to Tic22 was further demonstrated by immunological pull-down assays, in which Tic20 could be detected in an anti-Tic22 immunoprecipitate [21]. However, when solubilized chloroplast envelope membranes were subjected to sequential immunoaffinity chromatography, Tic20 and Tic22 could not be detected in the anti-Tic110-bound fraction after the samples had already been immunodepleted of Toc components (with anti-Toc34- and anti-Toc86/159-Sepharose). Additionally, in Blue Native PAGE (BN-PAGE) analysis of pea chloroplasts, Tic20 and Tic22 show a running behavior which is quite distinct from the Tic core complex as described by Küchler *et al.* [19] (containing Tic110, Tic62, and Tic55—see sections VIII and IX and Benz and Bölter, unpublished results). These results reveal a close connection of Tic20 to Tic22 and Toc as well as to translocating precursors, but also raise the question whether Tic20 and Tic110 are present in the same Tic complex or rather work independently of each other maybe in distinct Tic complexes.

In *Arabidopsis*, four homologues to Tic20 from pea (psTic20) seem to exist [12]. Two of these, namely atTic20-I (~62% identity) and atTic20-IV (~35% identity), are thought to represent real orthologues and/or paralogues of the pea gene, respectively. The other two (atTic20-II and atTic20-V) are more distantly related and less well conserved. Expression analysis of the former two genes in *Arabidopsis* has shown them to be the least abundant of all Tic components [57], but atTic20-I seems to be present in all tissues analyzed so far (including roots and etiolated seedlings [52]).

The importance of Tic20 for preprotein import and plant development has been studied using antisense lines in *Arabidopsis* [52]. Comparable to the previously described Tic components Tic110, Tic40 (and Hsp93-V), plants with reduced Tic20 expression showed a pale to chlorotic phenotype as well as altered chloroplast ultrastructure. Protein import experiments performed with preSSU indicated a specific defect in preprotein translocation across the IEM. Furthermore, it was shown that the accumulation of nuclear-encoded plastid proteins was impaired in antisense plants, whereas the abundance of major translocation components remained constant (Toc75, Toc159) or rather increased slightly, for example, for Tic110. These results implicate a significance of Tic20 in preprotein translocation, though they do not provide any evidence for a function as a channel protein.

## VI. Tic22: A Connection to Toc in the Intermembrane Space

As already mentioned, Tic22 was discovered along with Tic20 [21]. It is the only partially soluble Tic translocon component identified so far and is localized in the intermembrane space between the outer and inner envelope of chloroplasts. Tic22 has been detected by label-transfer cross-linking in direct contact to a precursor protein and seems to join the Tic complex only in the presence of Toc to form joint translocation sites [21].

A cyanobacterial homologue of Tic22 in *Synechocystis* (slr0924) has been identified and localized in the thylakoid lumen as well as the periplasmic space by immunogold labeling [58]. It has been found to comprise an essential protein which has not yet been determined for its counterpart in plants.

The import of Tic22 has been reported to follow a different pathway than the general import route, though the data are less than convincing [59]. It seems clear that Tic22 has a cleavable N-terminal transit peptide and needs only minor amounts of ATP to reach the intermembrane space. This is not surprising since the major amount of ATP necessary for import into the stroma is used by stromal chaperones [10, 20] (see section III) which may not play a role in import to the intermembrane space.

Becker *et al*. (2004) found that Tic22 is a member of a translocation-mediating intermembrane space complex consisting of Toc64, the J-protein Toc12, Hsp70, and Tic22 [60]. In the presence of ATP, Toc64 recruits Tic22 into the complex, where it directly interacts with Hsp70 and the translocating precursor protein. This scenario is reminiscent of the mitochondrial system where distinct complexes composed of small Tim proteins mediate translocation of proteins across the intermembrane space [36]. Thus, it cannot be excluded that an intermembrane space complex also exists in plastids that might be restricted to specific classes of proteins. This would be in line with the observation that Hsp70, Toc12, Toc64, and Tic22 are less abundant than the core components of Toc, namely Toc159, Toc34, and Toc75 [57].

Regarding the intermembrane space complex, a feature indicative of a redox-mediated import process comes to mind: the loop region of Toc12 following the HPD motif comprises two cysteine residues unique in the class of J-proteins. The disulfide (S–S) bridge formation between both cysteins was shown to stabilize the structure. This disulfide bond might also be important for the regulation of the import process itself, because reversible S–S bridge formation is a common mechanism to alter chloroplast enzyme activity [61]. It may be able to ''sense'' the redox state, the signal coming from the chloroplast through Tic32, 62, and 55, which will be discussed later (sections VII–IX).

## VII. Tic32: A Short Chain Dehydrogenase

Tic32 has been identified as an interaction partner of the N-terminal part of Tic110 [18]. The first 200 amino acids of Tic110 were heterologously expressed, purified on a nickel column and refolded to build a Tic110 affinity matrix which was then incubated with solubilized inner envelope proteins. Bound proteins were eluted from the matrix by a salt gradient resulting in a single protein being detached from Tic110-N. This protein of 32 kDa was then sequenced and identified as a member of the conserved class of short chain dehydrogenases (SDRs).

SDRs are a functionally heterogeneous protein superfamily found in all organisms and are defined by three distinct functional domains: an NAD(P)H-binding site at the N-terminus (TGXXXGXG), a β-sheet stabilizing motif in the central part (NNAG) and an active site containing an essential tyrosine residue at the C-proximal part [62]. SDRs are involved in as different functions as fatty acid and sugar metabolism, control of hormone ligand levels, transcriptional regulation, and apoptosis. Very often those enzymes are membrane bound and found in complexes with other membrane proteins [63, 64].

If inner envelope vesicles are treated with 1 M sodium chloride, pH 11.5 or 6 M urea and subsequently separated into pellet and supernatant, Tic32 is almost exclusively detected in the pellet fraction. This indicates that it represents an integral membrane protein, though secondary structure prediction programs place it within the class of soluble proteins [65]. Protease treatments of inner envelope vesicles showed that Tic32 is protected from degradation, indicating that it faces the stromal side of the IEM. It could be shown by chemical cross-linking and immunoprecipitations that Tic32 interacts with precursor proteins during the translocation process even at a very late stage of import. The affiliation of Tic32 to the inner envelope translocon was verified by coimmunoprecipitations. Tic32 not only interacts with Tic110 but is found in a complex with Tic40, Tic22, and Tic62. On analyzing T-DNA knockout mutants in *A. thaliana*, it became clear that Tic32 is essential for plant development. No homozygous plants were found in an extensive screen. Closer investigation of siliques from heterozygous plants revealed aborted seeds which contained aberrant embryos. This indicates that the presence of Tic32 is required for embryogenesis [18].

The tight interaction of the SDR Tic32 with the protein-conducting channel Tic110 reminds of the association of potassium channels with their regulative beta subunits [66]. Those Kvβ oxidoreductases couple the Kv channel activity with cellular redox regulation. Although it is still speculative that Tic32 is in fact responsible for the gating of the Tic110 channel, it could be shown that the heterologously expressed protein indeed

shows dehydrogenase activity and that the interaction with Tic110 is influenced by binding of NADPH (F. Hörmann and F. Chigri, personal communication).

Another fascinating feature of Tic32 was revealed on studying its import behavior. In contrast to almost all other components destined for an inner compartment of the chloroplast, Tic32 does not posses a cleavable N-terminal transit peptide, though the essential targeting information is contained within the 10 most N-terminal amino acids [67]. Not only the targeting of Tic32 differs from most other precursor proteins using the Toc/Tic machinery for the passage across the two envelope membranes, but also the actual route Tic32 takes into the plastids. Tic32 is imported in the presence of very low amounts of ATP (20 μM) at 4 °C and is not dependent on protease sensitive receptors on the chloroplast surface, as was demonstrated by import of Tic32 into thermolysin treated chloroplasts. It could also not be cross-linked to any Toc components indicating that Tic32 bypasses the known Toc machinery on an unknown pathway. The only member of the general import translocons found in connection with Tic32 was Tic110. But this does not necessarily point at an involvement of Tic110 in the import of Tic32 but could simply reflect the final situation of the two proteins being in very close proximity. Those results prove that beside the so-called general import pathway there must exist at least one other pathway into the chloroplast which circumvents the Toc/Tic machinery. Only one other example of an inner envelope protein being imported without a cleavable transit peptide has been published, although this protein (*c*hloroplast *e*nvelope *q*uinone *o*xidoreductase *h*omologue (ceQORH) does not contain any essential targeting information within the first 59 amino acids [68, 69].

## VIII. Tic62: The FNR-Binding Protein

Tic62 shares a specific feature with Tic32, which is a pyridine-nucleotide-binding site at the N-terminal part of the protein. It has been identified as a member of the Tic-core complex (containing Tic110, Tic55, and Tic62) isolated from BN-PAGE [19]. This result has been corroborated by coimmunoprecipitation, making Tic62 a *bona fide* member of Tic which coprecipitates together with Tic110, Tic40, Tic55, and Tic32. Besides containing an NADPH-binding site, which was experimentally confirmed, Tic62 exhibits a repetitive module of three repeats in the C-proximal part of the protein. It could be shown that this module interacts with a stromal enzyme, the ferredoxin oxidoreductase (FNR). FNR is known to be localized at the stromal face of thylakoids where it transfers electrons from ferredoxin to NADP, thereby playing a crucial role in photosynthesis and chloroplast

redox-controlled metabolism [70]. FNR is known to be released from thylakoids on oxidative stress [71], which might provide a pool of FNR free to interact with Tic62. Since transcription, translation as well as import of FNR itself is dependent on the chloroplast's redox state [72], association with Tic62 might reflect the presence of a signal cascade to sense the redox state of chloroplasts in which FNR could mediate electron transfer to Tic62-bound NADP.

## IX. Tic55: The Rieske-Family Member

Similar to Tic62, BN-PAGE was the method used to discover Tic55 [17]. In this case, Tic110 was used as a marker protein for a possible Tic complex. A band of ~280 kDa was shown to contain Tic110, ClpC, a 60-kDa protein (which we now know to be Tic62), a 55-kDa band and two others of 45 and 36 kDa, respectively. The 55-kDa protein was identified by peptide sequencing and screening of a pea cDNA library as a Rieske-type protein containing an additional mononuclear iron binding site. It was localized to the inner envelope membrane, shown to be resistant to treatment with pH 11.5 and to comprise a protease-protected fragment of 47 kDa. These results, together with secondary structure predictions, indicated that Tic55 spans the inner membrane twice, with the bulk of the protein being exposed to the stroma. It could be coimmunoprecipitated with Tic110, confirming the result from the BN-PAGE of it being a member of the inner envelope translocation machinery.

Rieske-type proteins play important roles in electron transfer, for example in the respiratory chain complexes. It has been shown that the chemical compound diethylpyrocarbonate (DEPC) inhibits the electron transfer activity of mitochondrial complex III (containing a Rieske-type cytochrome) by ethoxyformylation of a critical histidine residue [73]. The effect of DEPC on chloroplast protein import was studied using the precursor of SSU (preSSU). In the presence of 1 mM DEPC translocation of preSSU was drastically inhibited at the level of the inner envelope membrane, indicating that a Rieske-type protein like Tic55 plays a crucial role during protein translocation. Using a different experimental approach, direct binding of preSSU to Tic55 was demonstrated [17].

Database research revealed significant homology to bacterial aromatic ring-hydroxylating dioxygenases. Within the plant clade, close homologues were found in maize *l*ethal *l*eaf *s*pot (LLS1); today: *p*heophorbide *a o*xygenase (PAO) and *Arabidopsis*, classifying Tic55 as a member of CAO/PAO-like oxygenases [74]. Since the sequencing of the *Arabidopsis* genome has been completed [75], it became clear that Tic55 belongs to a tightly clustered family of plant oxygenases, containing LLS1 and Ptc52 (a chlorophyll a

oxygenase) which share a unique C-terminal part strictly conserved within this family comprising 25 members. Thus, those three proteins are closely related and likely to have derived from a common ancestor. The presence of PAO-like proteins correlates strongly with the emergence of oxygenic photosynthesis, indicating that oxygenase-like proteins may have evolved to allow oxygenic photosynthetic organisms to adapt and become fine tuned to oxygen/light levels in the water and on land [76]. The C-terminal conserved region of the PAO family comprises a cysteine, which supports Küchler's hypothesis of the redox-regulated Tic translocon; redox-regulated proteins exhibit conserved cysteine residues which could be targets of thioredoxin regulation.

Regarding the presence of Tic55 as a Rieske protein (see above) as well as Tic62 and Tic32 as possible dehydrogenases in the Tic machinery, the possibility of redox regulated import across the inner envelope becomes highly feasible. It was shown by Hirohashi and colleagues in 2001 that different ferredoxin and FNR isoforms are differentially imported dependent on the redox state of chloroplasts [77]. Those authors used light or darkness to alter the state of the plastidic NADPH pool. Küchler *et al.* used the chemical compounds deamino-NAD and HAR to demonstrate the influence of the chloroplasts' redox stage on import of those precursor proteins [19]. Also, it had already been reported by several researchers that the expression of nuclear-encoded plastid proteins is regulated by redox mechanisms [78]. Therefore, an additional redox controlled checkpoint at the level of protein import into the chloroplast seems only logical. Gene expression is in some cases a long-term response to the redox state [79], whereas import activity could represent a more instant response.

Taken together, we find three components in the Tic complex which exhibit features indicative of being involved in redox regulation, namely Tic55, Tic62, and Tic32. Although the *in vivo* substrates of the dehydrogenases are not known to date and the sequence of any electron transfer has to be pure speculation at this point, it seems highly probable that import across the inner chloroplast envelope is a redox-regulated process—either directly by influencing the gating behavior of the translocation pore Tic110 or indirectly by sensing and ''reporting'' the plastidic redox state.

## X. Traveling Back in Time

The evolutionary ancestors of plastids, the cyanobacteria, lack a protein import system. Thus, in the course of integrating the cyanobacterial endosymbiont into the host cell, a mechanism for protein import needed to be established along with or successively to the gene transfer from the organelle to the

nucleus. Only three of the Tic components have homologues in cyanobacteria, namely Tic55, Tic22, and Tic20 [53]. All other components seem to have derived from the eukaryotic host and/or underwent intensive mutations, which impedes their recognition. Reumann and Keegstra performed detailed phylogenetic studies of the Tic components [12]. According to their conclusions, two of the four Tic20 homologues in *Arabidopsis thaliana* correspond to the most ancient form (see section V). The other two homologues, including the pea orthologue, have acquired N-terminal targeting signals to exert their yet unknown function in chloroplast protein import.

The *Synechocystis* homologue of Tic22 (slr0924) has been studied experimentally and found to be localized in the thylakoid lumen as well as the periplasmic space (see section VI). Since slr0924 comprises an essential gene in *Synechocystis*, a possible role in photosynthesis or glucose metabolism was discussed, which would tally with the restriction of Tic22 proteins to cyanobacteria. Whether another *Arabidopsis* Tic22 homologue (atTic22-III) also resides in the chloroplast remains to be determined.

Interestingly, Tic20 and Tic22 also have homologues in the red algae *Porphyra purpurea* and *Cyanidium caldarium*, which are still encoded in the plastidic genome [12]. In *C. merolae*, one copy is still in the plastome, one other already found in the nucleus. Thus, looking at eukaryotes close to the base of the phylogenetic tree, representing intermediate stages of gene transfer, provides a neat possibility to get a direct glimpse of evolutionary processes.

Tic55-like proteins are found in cyanobacterial genomes as well [53]. Although no close homologues for Tic62 have been identified in prokaryotic organisms, just looking at the N-terminal part of the protein draws a different picture: the first 200 amino acids are closely related to cyanobacterial nucleoside-diphosphate sugar epimerases [12]. Thus, this Tic component seems to have been poached from a different function, extended by the C-terminal repeats and thereby integrated in the redox sensing of the chloroplast by binding to FNR. This seems to have happened rather late in evolution since the Tic62 orthologue in *Physcomitrella* also lacks the FNR-binding repeats (A. Stengel and B. Bölter, unpublished observation).

Generally, it seems quite likely that the subsequent assembly of translocon components originated from a primitive translocation machinery already present in the ancient ancestor of plastids, most likely mediating transport out of the cell. A very basic member of such an early translocon should be the protein conducting channel. This would suggest that it either derived from the cyanobacterial endosymbiont or was first in line to be targeted to the newly acquired organelle from the host organism. The fact that Tic20 has a cyanobacterial homologue would argue for Tic20 as the ancient channel, whereas the rather late transfer of the Tic20 gene to the

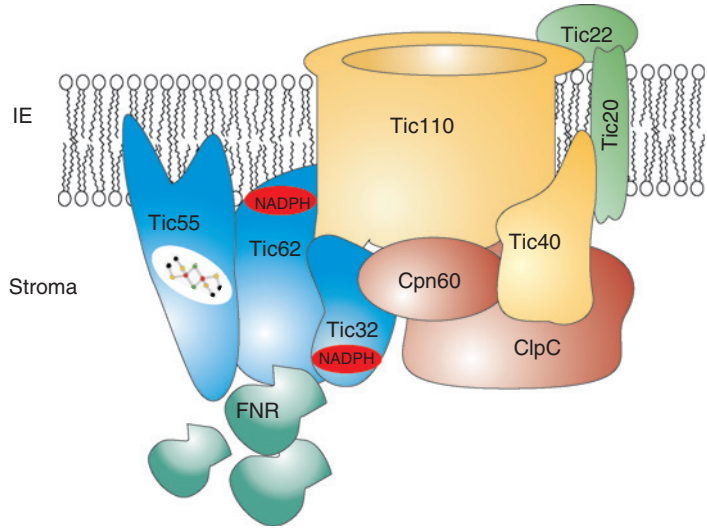



FIG. 17.1. Model of the Tic complex. Components of the Tic complex are designated as Tic, followed by their molecular masses in kDa, associated chaperones as Cpn60 and ClpC, respectively. The translocation channel Tic110 is associated with the cochaperone-like Tic40, the dehydrogenases Tic32 and Tic62, the Rieske-type Tic55 (all constituents involved in redox regulation are depicted in blue) and the chaperones ClpC/Cpn60. Tic22 has been shown to be involved in complex formation in the intermembrane space. For Tic20 a pore forming function is discussed.

nuclear genome speaks against this notion. It could also be imagined that until the recruitment of Tic20 and/or Tic110 as protein translocation channels a cyanobacterial Sec-component might have taken over this function. The principal possibility of retrograde transport has been shown for the eukaryotic Sec61 homologue [80].

Taken together, there are still a lot of questions to be solved concerning the Tic complex (a current model is depicted in Figure 17.1) but it became quite clear that the regulation of import across the inner membrane is different from that of the Toc machinery. While Toc is regulated via nucleotide exchange, a clearly eukaryotic feature, Tic seems to use the prokaryotic way of redox regulation for coordinating the needs of chloroplasts with the cellular metabolism.

## References


1. Margulis, L. (1970). *Origin of Eukaryotic cells*. Yale University Press, New Haven, CT, USA.
2. von Heijne, G. (1992). Membrane protein structure prediction: hydrophobicity analysis and the positive-inside rule. *J Mol Biol* 225:487–494.
3. Cline, K. (2000). Gateway to the chloroplast. *Nature* 403:148–149.
4. Schleiff, E., and Klosgen, R.B. (2001). Without a little help from 'my' friends: direct insertion of proteins into chloroplast membranes? *Biochim Biophys Acta* 1541:22–33.
5. Block, M.A., Dorne, A.J., Joyard, J., and Douce, R. (1983). Preparation and characterization of membrane fractions enriched in outer and inner envelope membranes from spinach chloroplasts. II. Biochemical characterization. *J Biol Chem* 258:13281–13286.
6. Cline, K., Andrews, J., Mersey, B., Newcomb, E.H., and Keegstra, K. (1981). Separation and characterization of inner and outer envelope membranes of pea chloroplasts. *Proc Natl Acad Sci USA* 78:3595–3599.
7. Werner-Washburne, M., Cline, K., and Keegstra, K. (1983). Analysis of Pea chloroplast inner and outer envelope membrane proteins by two-dimensional gel electrophoresis and their comparison with stromal proteins. *Plant Physiol* 73:569–575.
8. Davila-Aponte, J.A., Inoue, K., and Keegstra, K. (2003). Two chloroplastic protein translocation components, Tic110 and Toc75, are conserved in different plastid types from multiple plant species. *Plant Mol Biol* 51:175–181.
9. Inaba, T., Alvarez-Huerta, M., Li, M., Bauer, J., Ewers, C., Kessler, F., and Schnell, D.J. (2005). Arabidopsis tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* 17:1482–1496.
10. Kessler, F., and Blobel, G. (1996). Interaction of the protein import and folding machineries in theáchloroplast. *Proc Natl Acad Sci* 93:7684–7689.
11. Lubeck, J., Soll, J., Akita, M., Nielsen, E., and Keegstra, K. (1996). Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. *EMBO J* 15:4230–4238.
12. Reumann, S., Inoue, K., and Keegstra, K. (2005). Evolution of the general protein import pathway of plastids (review). *Mol Membr Biol* 22:73–86.
13. Schnell, D.J., Kessler, F., and Blobel, G. (1994). Isolation of components of the chloroplast protein import machinery. *Science* 266:1007–1012.
14. Wu, C., Seibert, F.S., and Ko, K. (1994). Identification of chloroplast envelope proteins in close physical proximity to a partially translocated chimeric precursor protein. *J Biol Chem* 269:32264–32271.
15. Vothknecht, U.C., and Soll, J. (2005). Chloroplast membrane transport: interplay of prokaryotic and eukaryotic traits. *Gene* 354:99–109.
16. Heins, L., Mehrle, A., Hemmler, R., Wagner, R., Kuchler, M., Hormann, F., Sveshnikov, D., and Soll, J. (2002). The preprotein conducting channel at the inner envelope membrane of plastids. *EMBO J* 21:2616–2625.
17. Caliebe, A., Grimm, R., Kaiser, G., Lubeck, J., Soll, J., and Heins, L. (1997). The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. *EMBO J* 16:7342–7350.
18. Hormann, F., Kuchler, M., Sveshnikov, D., Oppermann, U., Li, Y., and Soll, J. (2004). Tic32, an essential component in chloroplast biogenesis. *J Biol Chem* 279:34756–34762.
19. Kuchler, M., Decker, S., Hormann, F., Soll, J., and Heins, L. (2002). Protein import into chloroplasts involves redox-regulated proteins. *EMBO J* 21:6136–6145.

20. Akita, M., Nielsen, E., and Keegstra, K. (1997). Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. *J Cell Biol* 136:983–994.
21. Kouranov, A., Chen, X., Fuks, B., and Schnell, D.J. (1998). Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J Cell Biol* 143:991–1002.
22. Nielsen, E., Akita, M., Davila-Aponte, J., and Keegstra, K. (1997). Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO J* 16:935–946.
23. Bedard, J., and Jarvis, P. (2005). Recognition and envelope translocation of chloroplast preproteins. *J Exp Bot* 56:2287–2320.
24. Perry, S.E., and Keegstra, K. (1994). Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* 6:93–105.
25. Jackson, D.T., Froehlich, J.E., and Keegstra, K. (1998). The hydrophilic domain of Tic110, an inner envelope membrane component of the chloroplastic protein translocation apparatus, faces the stromal compartment. *J Biol Chem* 273:16583–16588.
26. Inaba, T., Li, M., Alvarez-Huerta, M., Kessler, F., and Schnell, D.J. (2003). AtTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *J Biol Chem* 278:38617–38627.
27. Grossman, A.R., Bartlett, S.G., Schmidt, G.W., and Chua, N.H. (1980). Post-translational uptake of cytoplasmically synthesized proteins by intact chloroplasts *in vitro*. *Ann N Y Acad Sci* 343:266–274.
28. Pain, D., and Blobel, G. (1987). Protein import into chloroplasts requires a chloroplast ATPase. *Proc Natl Acad Sci USA* 84:3288–3292.
29. Lubben, T.H., Donaldson, G.K., Viitanen, P.V., and Gatenby, A.A. (1989). Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell* 1:1223–1230.
30. Viitanen, P.V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G. H., Gatenby, A., and Soll, J. (1995). Functional characterization of the higher plant chloroplast chaperonins. *J Biol Chem* 270:18158–18164.
31. Lamppa, G.K., and Abad, M.S. (1987). Processing of a wheat light-harvesting chlorophyll a/b protein precursor by a soluble enzyme from higher plant chloroplasts. *J Cell Biol* 105:2641–2648.
32. Shanklin, J., DeWitt, N.D., and Flanagan, J.M. (1995). The stroma of higher plant plastids contain ClpP and ClpC, functional logs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. *Plant Cell* 7:1713–1722.
33. Halperin, T., Ostersetzer, O., and Adam, Z. (2001). ATP-dependent association between subunits of Clp protease in pea chloroplasts. *Planta* 213:614–619.
34. Andersson, F.I., Blakytny, R., Kirstein, J., Turgay, K., Bukau, B., Mogk, A., and Clarke, A. K. (2006). Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *J Biol Chem* 281:5468–5475.
35. Kovacheva, S., Bedard, J., Patel, R., Dudley, P., Twell, D., Rios, G., Koncz, C., and Jarvis, P. (2005). *In vivo* studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J* 41:412–428.
36. Rehling, P., Brandner, K., and Pfanner, N. (2004). Mitochondrial Import and the Twin-Pore Translocase. *Nat Rev Mol Cell Biol* 5:519–530.
37. Los Rios, P., Ben Zvi, A, Slutsky, O, Azem, A., and Goloubinoff, P. (2006). Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc Natl Acad Sci* 103:6166–6171.

38. Yadegari, R., Paiva, G., Laux, T., Koltunow, A.M., Apuya, N., Zimmerman, J.L., Fischer, R.L., Harada, J.J., and Goldberg, R.B. (1994). Cell differentiation and morphogenesis are uncoupled in arabidopsis raspberry embryos. *Plant Cell* 6:1713–1729.
39. Apuya, N.R., Yadegari, R., Fischer, R.L., Harada, J.J., Zimmerman, J.L., and Goldberg, R. B. (2001). The Arabidopsis embryo mutant schlepperless has a defect in the chaperonin-60alpha gene. *Plant Physiol* 126:717–730.
40. Apuya, N.R., Yadegari, R., Fischer, R.L., Harada, J.J., and Goldberg, R.B. (2002). RASPBERRY3 gene encodes a novel protein important for embryo development. *Plant Physiol* 129:691–705.
41. Ko, K., Budd, D., Wu, C., Seibert, F., Kourtz, L., and Ko, Z.W. (1995). Isolation and characterization of a cDNA clone encoding a member of the Com44/Cim44 envelope components of the chloroplast protein import apparatus. *J Biol Chem* 270:28601–28608.
42. Stahl, T., Glockmann, C., Soll, J., and Heins, L. (1999). Tic40, a new ''old'' subunit of the chloroplast protein import translocon *J Biol Chem* 274:37467–37472.
43. Chou, M.L., Fitzpatrick, L.M., Tu, S.L., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J.Z., Keegstra, K., and Li, H.M. (2003). Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J* 22:2970–2980.
44. Frydman, J., and Hohfeld, J. (1997). Chaperones get in touch: the Hip-Hop connection. *Trends Biochem Sci* 22:87–92.
45. Hohfeld, J., Minami, Y., and Hartl, F.U. (1995). Hip, a novel cochaperone involved in the eukaryotic Hsc70/Hsp40 reaction cycle. *Cell* 83:589–598.
46. Webb, M.A., Cavaletto, J.M., Klanrit, P., and Thompson, G.A. (2001). Orthologs in *Arabidopsis thaliana* of the Hsp70 interacting protein Hip. *Cell Stress Chaperones* 6:247–255.
47. Buchner, J. (1999). Hsp90 & Co.: a holding for folding. *Trends Biochem Sci* 24:136–141.
48. Richter, K., Muschler, P., Hainzl, O., Reinstein, J., and Buchner, J. (2003). Sti1 is a non-competitive inhibitor of the Hsp90 ATPase. Binding prevents the N-terminal dimerization reaction during the atpase cycle. *J Biol Chem* 278:10328–10333.
49. Abbas-Terki, T., Donze, O., Briand, P.A., and Picard, D. (2001). Hsp104 interacts with Hsp90 cochaperones in respiring yeast. *Mol Cell Biol* 21:7569–7575.
50. Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21:932–939.
51. Niwa, Y., Kato, T., Tabata, S., Seki, M., Kobayashi, M., Shinozaki, K., and Moriyasu, Y. (2004). Disposal of chloroplasts with abnormal function into the vacuole in *Arabidopsis thaliana* cotyledon cells. *Protoplasma* 223:229–232.
52. Chen, X., Smith, M.D., Fitzpatrick, L., and Schnell, D.J. (2002). *In vivo* analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* 14:641–654.
53. Reumann, S., and Keegstra, K. (1999). The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. *Trends Plant Sci* 4:302–307.
54. Ma, Y., Kouranov, A., LaSala, S.E., and Schnell, D.J. (1996). Two components of the chloroplast protein import apparatus, IAP86 and IAP75, interact with the transit sequence during the recognition and translocation of precursor proteins at the outer envelope. *J Cell Biol* 134:315–327.
55. Rassow, J., Dekker, P.J., van Wilpe, S., Meijer, M., and Soll, J. (1999). The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J Mol Biol* 286:105–120.
56. Kouranov, A., and Schnell, D.J. (1997). Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. *J Cell Biol* 139:1677–1685.

57. Vojta, A., Alavi, M., Becker, T., Hormann, F., Kuchler, M., Soll, J., Thomson, R., and Schleiff, E. (2004). The protein translocon of the plastid envelopes. *J Biol Chem* 279:21401–21405.
58. Fulda, S., Norling, B., Schoor, A., and Hagemann, M. (2002). The Slr0924 protein of *Synechocystis* sp. *strain* PCC 6803 resembles a subunit of the chloroplast protein import complex and is mainly localized in the thylakoid lumen. *Plant Mol Biol* 49:107–118.
59. Kouranov, A., Wang, H., and Schnell, D.J. (1999). Tic22 is targeted to the intermembrane space of chloroplasts by a novel pathway. *J Biol Chem* 274:25181–25186.
60. Becker, T., Hritz, J., Vogel, M., Caliebe, A., Bukau, B., Soll, J., and Schleiff, E. (2004). Toc12, a novel subunit of the intermembrane space preprotein translocon of chloroplasts. *Mol Biol Cell* 15:5130–5144.
61. Pfannschmidt, T. (2003). Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci* 8:33–41.
62. Kallberg, Y., Oppermann, U., Jornvall, H., and Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). *Eur J Biochem* 269:4409–4417.
63. Simon, A., Romert, A., Gustafson, A.L., McCaffery, J.M., and Eriksson, U. (1999). Intracellular localization and membrane topology of 11-cis retinol dehydrogenase in the retinal pigment epithelium suggest a compartmentalized synthesis of 11-cis retinaldehyde. *J Cell Sci* 112:549–558.
64. Simon, A., Hellman, U., Wernstedt, C., and Eriksson, U. (1995). The retinal pigment epithelial-specific 11-cis retinol dehydrogenase belongs to the family of short chain alcohol dehydrogenases. *J Biol Chem* 270:1107–1112.
65. Tusnady, G.E., and Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* 283:489–506.
66. Bahring, R., Milligan, C.J., Vardanyan, V., Engeland, B., Young, B.A., Dannenberg, J., Waldschutz, R., Edwards, J.P., Wray, D., and Pongs, O. (2001). Coupling of voltage-dependent potassium channel inactivation and oxidoreductase active site of Kvbeta subunits. *J Biol Chem* 276:22923–22929.
67. Nada, A., and Soll, J. (2004). Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *J Cell Sci* 117:3975–3982.
68. Kleffmann, T., Hirsch-Hoffmann, M., Gruissem, W., and Baginsky, S. (2006). plprot: a comprehensive proteome database for different plastid types. *Plant Cell Physiol* 47:432–436.
69. Miras, S., Salvi, D., Ferro, M., Grunwald, D., Garin, J., Joyard, J., and Rolland, N. (2002). Non-canonical transit peptide for import into the chloroplast. *J Biol Chem* 277:47770–47778.
70. Arakaki, A.K., Ceccarelli, E.A., and Carrillo, N. (1997). Plant-type ferredoxin-NADP + reductases: a basal structural framework and a multiplicity of functions. *FASEB J* 11:133–140.
71. Palatnik, J.F., Valle, E.M., and Carrillo, N. (1997). Oxidative stress causes ferredoxin-NADP + reductase solubilization from the thylakoid membranes in methyl viologen-treated plants. *Plant Physiol* 115:1721–1727.
72. Pfannschmidt, T., Schutze, K., Brost, M., and Oelmuller, R. (2001). A novel mechanism of nuclear photosynthesis gene regulation by redox signals from the chloroplast during photosystem stoichiometry adjustment. *J Biol Chem* 276:36125–36130.
73. Ohnishi, T., Meinhardt, S.W., von Jagow, G., Yagi, T., and Hatefi, Y. (1994). Effect of ethoxyformic anhydride on the Rieske iron-sulfur protein of bovine heart ubiquinol: cytochrome *c* oxidoreductase. *FEBS Lett* 353:103–107.

74. Gray, J., Wardzala, E., Yang, M., Reinbothe, S., Haller, S., and Pauli, F. (2004). A small family of LLS1-related non-heme oxygenases in plants with an origin amongst oxygenic photosynthesizers. *Plant Mol Biol* 54:39–54.
75. The Arabidopsis Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815.
76. Ting, C.S., Rocap, G., King, J., and Chisholm, S.W. (2002). Cyanobacterial photosynthesis in the oceans: the origins and significance of divergent light-harvesting strategies. *Trends Microbiol* 10:134–142.
77. Hirohashi, T., Hase, T., and Nakai, M. (2001). Maize non-photosynthetic ferredoxin precursor is mis-sorted to the intermembrane space of chloroplasts in the presence of light. *Plant Physiol* 125:2154–2163.
78. Oswald, O., Martin, T., Dominy, P.J., and Graham, I.A. (2001). Plastid redox state and sugars: interactive regulators of nuclear-encoded photosynthetic gene expression. *Proc Natl Acad Sci USA* 98:2047–2052.
79. Pfannschmidt, T., Nilsson, A., and Allen, J.F. (1999). Photosynthetic control of chloroplast gene expression. *Nature* 397:625–628.
80. Schnell, D.J., and Hebert, D.N. (2003). Protein translocons: multifunctional mediators of protein translocation across membranes. *Cell* 112:491–505.

# 18

# *The Sec and Tat Protein Translocation Pathways in Chloroplasts*

KENNETH CLINE[a] • STEVEN M. THEG[b]

[a]*Horticultural Sciences Department and Plant Molecular and Cellular Biology*
*University of Florida*
*Gainesville, FL 32611, USA*

[b]*Section of Plant Biology*
*University of California at Davis*
*Davis, CA 95616, USA*

## I. Abstract

Most chloroplast proteins are encoded by nuclear genes, translated in the cytosol as precursor proteins, and posttranslationally imported into the chloroplast. A subset of imported proteins is further localized to the thylakoid membrane and lumen by mechanisms conserved from the cyanobacterial endosymbiont that evolved into the chloroplast. The Sec and *T*win *a*rginine *t*ranslocation (Tat) pathways are the major systems for transporting proteins across the thylakoid membrane into the lumen. Both systems employ hydrophobic cleavable signal peptides for targeting, but Tat signal peptides also contain an essential twin arginine motif. Biochemical studies indicate that the thylakoid Sec system operates similarly to the *Escherichia coli* Sec system, that is a chloroplast SecA powers transport of unfolded protein substrates through a fixed cpSecYE channel. Indirect evidence also suggests that the thylakoid Sec system can integrate plastid-encoded



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25018-8

multispanning membrane proteins cotranslationally. The Tat pathway is a newly discovered translocation system that can transport folded protein domains using the $\Delta\mu^{H^+}$ as the sole energy source. Three membrane proteins, *H*igh *c*hlorophyll *f*luorescence 106 (Hcf106), *T*hylakoid *a*ssembly 4 (Tha4), and cpTatC constitute the components of the Tat machinery. Precursor proteins bind to a large cpTatC–Hcf106 complex by contact of the signal peptide twin arginine region to cpTatC and its hydrophobic core to Hcf106. This triggers recruitment of a Tha4 oligomer, setting the stage for transport. During the translocation step, the Tha4 oligomer undergoes a conformation shift that aligns its amphipathic helices and carboxyl tails, possibly in association with the bilayer interface. These results have been interpreted in a general model in which the Tha4 oligomer facilitates passage of the substrate across the lipid bilayer.

## II. Overview of Protein Trafficking to the Plant Thylakoid Membrane and Lumen

Chloroplasts are structurally complex organelles with three membranes and three aqueous compartments. A double-envelope membrane system encloses the aqueous stroma; the thylakoid membrane, embedded in the stroma, encloses the aqueous lumen, which is the innermost compartment of the chloroplast. Chloroplasts contain 2000–3000 different proteins [1], most of which are stromal proteins and most of which are encoded in the cell nucleus. The thylakoid membrane and lumen each contain ~100 different proteins [2, 3]. About 50% of thylakoid membrane proteins are encoded by plastid genes [4, 5] and translated on 70S ribosomes. The remaining thylakoid membrane proteins and all of the known lumenal proteins are encoded on nuclear genes, synthesized in the cytosol as precursor proteins, and posttranslationally imported into the chloroplast [6, 7]. Because of robust *in vitro* assays, much more is known about the targeting pathways of the nuclear-encoded thylakoid proteins. Chloroplasts are readily isolated from a number of plant species and are active in importing precursor proteins that are either made by *in vitro* translation or produced by heterologous expression in bacteria [8]. Studies with chloroplast import assays have shown that localization of nuclear-encoded thylakoid proteins is a two-step process [6, 7], wherein precursor proteins are first imported across the chloroplast envelope and transiently appear in the stroma as soluble intermediates before translocation into thylakoids [9, 10]. Import into the chloroplast is directed by stroma targeting transit peptides that are present on the N-terminus of precursors. Cleavage of transit peptides by the stromal processing peptidase exposes thylakoid-targeting signals that direct

integration into or transport across the thylakoid membrane (shown for the Sec and Tat pathways in Figure 18.1).

The pathways and mechanisms for thylakoid targeting have been examined in assays either with intact chloroplasts or with isolated thylakoids. There are several ways of manipulating chloroplast import assays to specifically inhibit the thylakoid translocation step. These include dissipating the proton gradient with ionophores [11, 12], depleting internal nucleoside triphosphates (NTPs) with glycerate [13], inhibiting SecA with azide [14, 15], and competing for precursor transport with saturating quantities of other precursors [16]. However, the isolated thylakoid translocation assay allows more precise control of experimental conditions and has been the

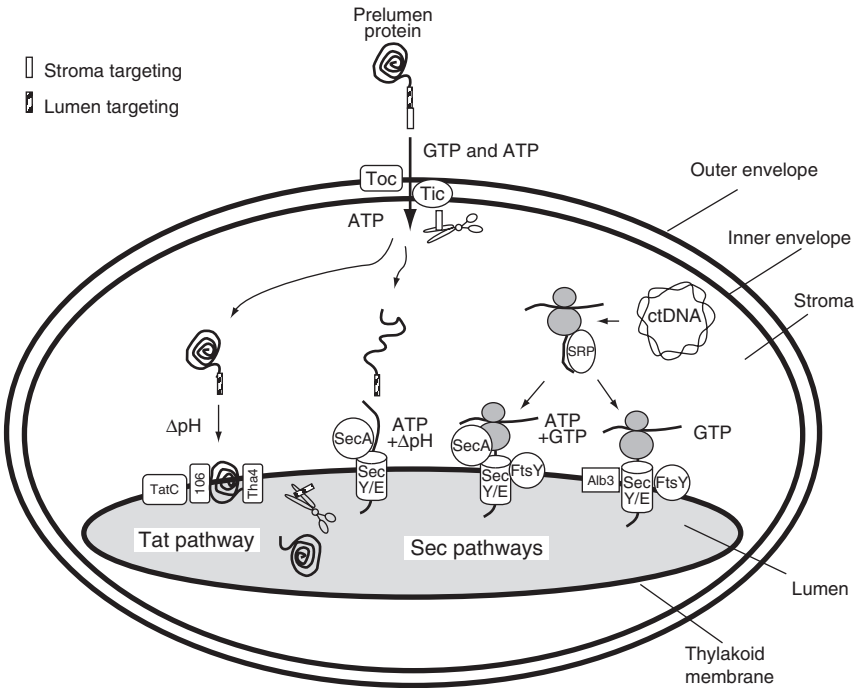



FIG. 18.1. Model for Sec and Tat protein transport pathways in chloroplasts. Nuclear-encoded thylakoid lumenal proteins contain bipartite-targeting signals. A stromal-targeting transit peptide directs import into the stroma by the Toc and Tic import complexes in the envelope. Cleavage of the transit peptide exposes signal peptides that specifically target one subgroup of precursors to the Tat pathway, which transports proteins in folded conformation, and another subgroup to the cpSecA-cpSecYE pathway, which transports proteins in unfolded conformation. Some chloroplast-encoded thylakoid proteins appear to be integrated by ribosome or ribosome-cpSecA coupled to the cpSecYE channel (see text).

assay of choice for mechanistic studies of the translocases (http://www.hos.ufl.edu/clineweb/).

Biochemical characterization of thylakoid protein transport in the early 1990s determined that there were two precursor-specific pathways for lumenal protein transport that could be distinguished by their functional requirements [12, 17] and could also be selectively competed with saturating quantities of precursors [16]. One of these pathways was found to be homologous to the well-characterized bacterial Sec system [18, 19]. The other pathway turned out to be a novel translocation system. It was initially called the ΔpH-dependent pathway because of its requirement for the thylakoidal ΔpH as sole energy source. It has been renamed "Tat" for *T*win *a*rginine *t*ranslocation to reflect the presence of two contiguous and essential arginine residues in the signal peptides of precursors that use this pathway. It is now known that Tat systems are widely represented in prokaryotes and prokaryote-derived organelles. Surprisingly, they are, unlike almost all other translocation systems, able to transport fully folded proteins.

There are four known pathways for protein localization to thylakoids [6, 7, 20]. In addition to Sec and Tat, the cpSRP and "Spontaneous" pathways are involved in membrane protein integration and are covered elsewhere in this volume. All of these pathways are highly homologous to protein transport and integration pathways found in prokaryotes and assuredly are conserved from the cyanobacterial endosymbiont that gave rise to the chloroplast. This chapter will focus on the chloroplast Sec and Tat protein transport pathways (Figure 18.1).

## III. Targeting to the Sec and Tat Pathways

Most of the known substrates of the thylakoid Sec and Tat pathways are resident proteins of the thylakoid lumen. In addition, simple membrane proteins with large lumenal domains, for example cytochrome *f* (Cytf) and photosystem 1 F subunit (PsaF), are integrated by the Sec pathway (see later) and two single span membrane proteins are integrated by the Tat pathway [21, 22]. Thylakoid Sec may also integrate multispanning membrane proteins such as D1 (see later). Virtually all of the lumenal and single-span membrane proteins are targeted to their respective pathways by cleavable hydrophobic signal peptides with features similar to bacterial signal peptides (Figure 18.2). These include a charged N-terminal region (N), a hydrophobic core region (H), and polar C-terminal region (C) with an A-X-A motif for cleavage by the thylakoid processing peptidase (TPP). Characteristics of thylakoid Sec and Tat signal peptides have been reviewed [23–25]. One notable feature is that both thylakoid Sec and Tat signal

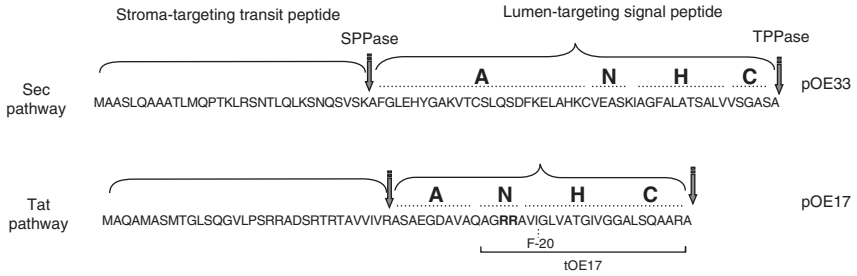



FIG. 18.2. Domain structure of targeting peptides for Sec- and Tat-directed thylakoid lumenal proteins. The stroma-targeting transit peptides are cleaved by the stromal processing peptidase (SPP), exposing the lumen-targeting signal peptides. Lumen targeting signal peptides are cleaved after thylakoid transport by the TPP. The tOE17 Tat pathway precursor is N-terminally truncated to remove the stroma-targeting peptide and the nonessential A domain. The essential twin arginine motif is in the larger bold font. The substitution of phenylalanine (F) for valine (V) at the −20 position of tOE17 alters the strength of its binding to cpTatC–Hcf106 (see text).

peptides generally possess unusually long N domains that usually contain acidic residues near the N-terminus. We previously referred to these acidic regions as ''A'' domains [23]. A domains can be deleted from certain precursors without adversely affecting transport competence *in vitro*. Deleting the A domains of 23-kD subunit of the oxygen evolving complex (OE23) and 17-kD subunit of the oxygen evolving complex (OE17) precursors actually improves Tat transport competence [26] and increases OE17-binding affinity for Tat components [27]. The physiological relevance of A domains is unclear; presumably they regulate access of precursors to the translocases.

A twin arginine (RR) motif in the N-domain of signal peptides is a specific targeting signal for the Tat pathway. Nearly all known Tat pathway precursors contain RR in their signal peptides. One exception is the Rieske FeS protein that has a KR in an uncleaved signal peptide [21]. Mutating the twin arginine motif to QQ or even KK prevents translocation as well as several other measurable interactions with Tat pathway machinery [26, 28, 29] (and see later). However, such mutated precursors do seem to interact with the Tat pathway because they stimulate proton counterflow [30] and can alter the transport behavior of authentic precursor proteins [30]. Based on the presence or absence of the twin arginines, prediction programs suggest that about 50% of lumenal proteins use the Tat pathway and 50% the Sec pathway [2].

Other features of Tat signal peptides make them incompatible with the Sec pathway. For example, the frequent occurrence of basic residues in Tat signal peptide C domains [31] and a somewhat reduced hydrophobicity of

Tat signal peptide H domains [2] appear to prevent transport of Tat pathway precursors by the Sec pathway. In this regard, dual targeting signal peptides have been engineered by fusing a twin arginine N domain to the H and C domains of a Sec pathway signal peptide [26]. There also appears to be a general lack of compatibility of Tat pathway mature proteins with the Sec machinery [26, 32], as chimeric proteins with a Sec signal peptide and a Tat pathway mature protein are not transported or even recognized by the Sec machinery. This incompatibility may relate to tight folding of Tat precursors prior to transport (see section V.B.1).

## IV. The Sec Transport System in Chloroplasts

### A. Introduction

Sec systems are evolutionarily conserved protein translocation machineries that are present in the eukaryotic endoplasmic reticulum [33], the archaeal plasma membrane [34], the eubacterial plasma membrane [35–37], and thylakoid membranes of plant and algal chloroplasts. Sec systems are versatile translocases capable of transporting soluble protein completely across the bilayer, of integrating multispanning membrane proteins, and a combination of both for those proteins that contain large transported domains as well as transmembrane domains. A common mechanistic feature of all Sec systems is that the substrates are transported in an unfolded conformation through a protein-conducting channel. The Sec channel consists of two essential proteins, the multispanning SecY (Sec61$\alpha$) and the usually single spanning SecE (Sec61$\gamma$). Sec channel complexes usually also contain a nonessential additional subunit (SecG in bacteria). Protein transport and integration is accomplished by a variety of accessory factors that include molecular chaperones on both sides of the membrane, targeting chaperones and their receptors, translocation motors, and membrane protein assembly chaperones. The specific accessory factors depend on the evolutionary origins of the organism and membrane system.

The thylakoid Sec system derives from the cyanobacterial endosymbiont that evolved into the chloroplast and thus is most closely related to eubacterial Sec systems. Among these, the *Escherichia coli* Sec system is the best characterized. *E. coli* Sec operates in a modular fashion. The SecYEG channel couples with the SecA ATPase for transporting soluble periplasmic proteins and with the ribosome for integrating multispanning membrane proteins. SecA binds both the signal peptide and mature domain of precursors and SecYEG. Through a cycle of ATP binding, hydrolysis, and release, SecA undergoes major conformational changes that serve to thread 20–30 residue segments of preproteins through the channel. The ribosome is thought to

perform a similar function via GTP hydrolysis during the peptide bond formation [38]. However, the ribosome apparently does not provide an equivalent motive force because proteins with large periplasmic loops require both the ribosome and SecA for translocation [36]. *E. coli* possesses other accessory factors for transport of subgroups of proteins that include the cytoplasmic chaperone SecB [39], the heterotrimeric membrane complex SecDFyajC [35], and the signal recognition particle (SRP) and its receptor FtsY that target nascent chain ribosomal complexes to SecYEG [33]. In addition, during insertion of membrane proteins, the membrane protein YidC associates with SecYEG and may chaperone transmembrane segments as they partition from the SecYEG channel into the lipid bilayer [40].

## B. Thylakoid Components

As previously discussed [20, 24], chloroplasts possess some of these components, but lack others. Chloroplasts possess SecA (cpSecA), SecY (cpSecY), and SecE (cpSecE), but lack SecB, SecG, and SecDFyajC. Chloroplasts have an SRP (cpSRP) and an FtsY (cpFtsY) [38], but the SRP lacks an RNA component and instead has a novel protein subunit called cpSRP43. Chloroplasts also contain a YidC homologue. The *Arabidopsis* genome encodes two chloroplast SecY family members (cpSecY and cpSecY2), two SecA family members (cpSecA and cpSecA2), and two YidC homologous family members (Alb3 and Alb4) [41]. Current knowledge regarding the operation of the thylakoid Sec system pertains to one of each family member, that is cpSecY, cpSecA, and Alb3.

## C. Capabilities and Operation of the Thylakoid Sec System

### 1. *Sec Transport of Lumenal Proteins in the cpSecA-cpSecYE Configuration*

The thylakoid Sec system in the cpSecA-cpSecYE configuration for transporting soluble protein precursors has been examined in most detail. *In vitro* analyses of transport are conducted with washed thylakoids, purified cpSecA [18], or stromal extract as a source of cpSecA, and *in vitro* translated precursor proteins. Potentially necessary chaperones are apparently provided by the stromal or translation extracts. Studies with this system indicate that thylakoid cpSecA/cpSecYE operates similarly to the *E. coli* system. Precursors bind to thylakoids in the absence of ATP (presence of apyrase) in a reaction that is stimulated by added cpSecA [42, 43]. Chemical cross-linking of membrane-bound precursors stabilizes a complex containing precursor, cpSecA, and cpSecY [42, 43]. Bound precursors can be chased into the thylakoid lumen with ATP without additional cpSecA [43]. The translocation

step is inhibited by the SecA inhibitor azide [15, 18] or by pretreating membranes with anti-cpSecY IgGs [44], is stimulated by the thylakoidal ΔpH [17, 45], and requires that the precursor can be unfolded. For example, a chimeric precursor containing dihydrofolate reductase (DHFR) in a methotrexate-stabilized folded conformation caused translocation arrest of the thylakoid Sec system [46, 47]. Time course analysis of the progression of precursor across the thylakoid Sec pathway demonstrated that the N-terminal signal peptide reached the lumenal signal peptidase while the bulk of the protein was accessible from the stromal (*cis*) side of the membrane, suggesting a linear mode of translocation [48].

### 2. *Cotranslational Integration of Membrane Proteins by Sec in the Ribosome-cpSecYE Configuration*

Thylakoids possess a substantial number of nuclear encoded and plastid-encoded multispanning membrane proteins that are potential candidates for Sec-mediated integration. Unfortunately, it has not been possible to use the isolated thylakoid assay to determine the integration pathway of most of these proteins. For example, the plastid encodes many multispanning membrane proteins. Because these proteins are frequently ligated to pigments and other cofactors and are assembled deep within core photosystem complexes, their assembly has been difficult to reconstitute with isolated thylakoids. In addition, their integration is cotranslational and must be reconstituted with an organelle-free plastid translation system programmed with exogenous RNA, and this system is technically difficult. Thus, indirect evidence has been used to infer pathway usage as described for the following two plastid-encoded proteins Cytf and D1.

Cytf consists of a large lumenal domain and a single membrane anchor. The requirement of cpSecA for Cytf integration was indicated by its accumulation as preCytf in the maize cpSecA-null mutant called *tha1* [49]. The involvement of cpSecA in Cytf integration was confirmed by *in vitro* reconstituted assays [50, 51]. The fact that Cytf is translated on thylakoid-bound ribosomes implies that Cytf is integrated by a cpSecA and ribosome-coupled cpSecYE. Cytf may be targeted to the cpSec channel by the chloroplast SRP because the Cytf signal peptide was found to cross-link to the 54-kDa subunit of the chloroplast SRP (cpSRP54) in a wheat germ translation system [52]. However, this point requires further study because in a chloroplast translation system cross-links of preCytf to cpSRP54 were not observed [51].

The multispanning photosystem 2 A subunit (PsbA) protein called D1 is also cotranslationally integrated into thylakoids. The translation and maturation of D1 is readily reconstituted *in organello* with intact chloroplasts.

Analysis of D1 ribosome nascent chain complexes produced during *in organello* protein synthesis showed that cpSecY is associated with nascent D1 chains, but not full-length D1, and is also bound to ribosomes in a salt-sensitive and puromycin-insensitive fashion [53]. D1 nascent chains produced by an organelle-free translation system can be cross-linked to cpSRP54 [54]. Taken together, these findings suggest that D1 ribosome nascent chain complexes are targeted to the cpSecYE complex by the chloroplast SRP, where D1 is integrated by the ribosome-coupled cpSecYE configuration. Alb3 appears necessary for efficient assembly of D1 into PS II reaction centers in *C. reinhardtii*, suggesting that Alb3 associates with cpSecYE for membrane protein assembly [55].

The existence of cpSecY in several different membrane complexes supports the idea that thylakoid Sec operates in multiple configurations. Blue native polyacrylamide gel electrophoresis (BN-PAGE) analysis of dodecylmaltoside-solubilized spinach thylakoids detected cpSecY in an ~100-kDa complex and in another complex near the top of the gel that comigrated with ribosomes [53]. Other studies have identified a complex that contains cpSecY, cpSecE, and Alb3 [56]. Thus, it appears from these studies that the thylakoid Sec system operates in the same configurations as the *E. coli* Sec system, with the cpSecYE channel playing the central and indispensable role. The relative severity of maize mutants lacking cpSecA versus cpSecY supports this notion. The maize cpSecA null strain *tha1* is reduced in many, but not all, thylakoid proteins and protein complexes [49]. The maize cpSecY null strain, although possessing plastids and surviving to the seedling stage, is virtually devoid of thylakoid membranes [57].

### D. cpSecY2 and cpSecA2

cpSecY is localized exclusively to thylakoids in mature chloroplasts [58], and its function appears limited to thylakoids [57]. The *Arabidopsis* and rice genomes encode a second predicted chloroplast SecY (cpSecY2) that is significantly diverged on the sequence level from cpSecY. cpSecY2 is also widely represented in the EST collections of a range of plant species. cpSecY2 appears to play a different role in plastid biogenesis because an *Arabidopsis* cpSecY2 null line is lethal at the globular embryo stage in developing seeds and thus does not even progress to the seedling stage (D. Fernandez, personal communication, Madison, WI). Promoter-swapping experiments between cpSecY and cpSecY2 demonstrate that the different phenotypes of the null mutants are not due to different expression patterns, but rather to different functions of the proteins (C. Skalitzky and D. Fernandez, personal communication, Madison, WI). These findings suggest that cpSecY2 is essential for biogenesis of a membrane other than

the thylakoids. cpSecY2 has a predicted chloroplast-targeting signal and *in vitro* translated cpSecY2 was imported into pea chloroplasts in an *in vitro* import assay (J. Martin and K. Cline, unpublished data). The imported protein was recovered in both the envelope and the thylakoids subfractions (J. Martin and K. Cline, unpublished data). Thus, one possibility is that cpSecY2 participates in translocation of inner plastid envelope proteins and/or interenvelope space proteins. In that regard, a study of an *Arabidopsis* signal peptidase 1 protein reported a dual localization in chloroplast envelope and in thylakoids [59], and cpSecY of the alga *Cyanophora paradoxa* is reported to be present in envelope and thylakoids [60]. A second SecA homologue with a predicted chloroplast-targeting signal is encoded in the *Arabidopsis* genome. This putative cpSecA2 may be a functional partner of cpSecY2, but is as yet uncharacterized.

### E. Prospects

Future progress in identifying substrates of cpSecY and cpSecY2 will likely require production and analysis of conditional mutants, depletion strains, or with other protein-disabling methodologies such as CALI [61]. Plastid-encoded thylakoid membrane proteins are the probable substrates for ribosome-coupled cpSecYE. On the other hand, nuclear-encoded thylakoid or envelope multispanning membrane proteins might involve an unknown configuration of cpSecYE as they are imported into the chloroplast after release from 80S ribosomes. Mechanistic characterization of cpSecY-mediated integration of plastid-encoded membrane proteins may come from use of the organelle-free chloroplast translation system and *in vitro* coupled integration assay [51], possibly with simpler substrates as an inroad to approaching the more highly complex proteins that make up photosynthetic complexes. Mechanistic studies of the nuclear-encoded substrates of cpSecY and cpSecY2 may provide novel insights into the ways chloroplasts have adapted to the posttranslational integration imposed as a result of endosymbiosis.

## V. The Tat System in Chloroplasts

### A. Introduction

Studies of the thylakoid Sec system benefited from the extensive knowledge of Sec systems in other organisms. By contrast, characterization of Tat systems began with thylakoid studies in the early 1990s and the first Tat component, maize *H*igh *c*hlorophyll *f*luorescence 106 (Hcf106), was

identified in 1997 [62]. Thus, the Tat system is being approached simultaneously in a variety of organisms. The presence of Hcf106 homologous genes and twin arginine signal peptides in sequenced organisms indicates that the Tat system is widely represented among prokaryotes and prokaryote-derived organelles, but absent from fungi and animals [63, 64]. Tat systems are uniformly present in the chloroplasts of plants and algae. In addition, one Tat component, TatC, is usually encoded in plant and algal mitochondrial genomes, although the presence of a functional Tat system in mitochondria has not been demonstrated. In eubacteria, Tat substrates represent a small subset of the total secretome (less than 10% in *E. coli*) and are frequently metal cofactor-containing proteins that employ cytosolic mechanisms for cofactor insertion. Several studies have shown that bacterial Tat substrates, which are as large as 7 nm [65], are folded prior to transport [66, 67]. Of the ~50 predicted substrates of the thylakoid Tat system, only a few possess cofactors. One possible reason for the continued maintenance of this system in chloroplasts is that imported Tat substrates are rapidly folded by stromal chaperones and may be difficult to unfold by the thylakoid Sec system.

Mechanistic investigations of Tat systems are being done primarily with bacterial and plant thylakoid systems. Work in *E. coli* has been particularly effective in examining the *in vivo* activity of the system and structural characteristics of Tat complexes. Plant thylakoids permit a robust *in vitro* assay, well-characterized energetics, and amenable biochemistry. The following discussion will summarize results directed toward a mechanistic understanding of the thylakoid system, with reference to similarities and differences between thylakoids and bacterial Tat systems.

## B. Capabilities and Requirements of the Thylakoid Tat System

### *1. Folded and Unfolded Proteins*

The thylakoid Tat system appears capable of transporting both tightly folded and unfolded proteins. The observation that at least two natural Tat substrates adopt tight folds suggested folded protein transport [68, 69]. This was directly demonstrated in experiments where internally cross-linked bovine pancreatic trypsin inhibitor (BPTI) was efficiently transported when fused to the C-terminus of precursor to OE17 (pOE17) [70], and in which DHFR locked in a folded conformation was efficiently transported when fused to the C-terminus of precursor to OE23 (pOE23) [46]. An experiment, showing that thylakoid Tat, but not Sec, can transport GFP, is also indicative of folded protein transport [71]. Predicted thylakoid Tat substrates range in size from ~2 kDa to over 60 kDa, although the folding

status of nearly all of these substrates is unknown [2]. Thus, it is not clear if thylakoid Tat can transport the very large folded protein domains that the *E. coli* Tat system appears to transport [72].

A number of investigators have asked whether the thylakoid system can transport unfolded proteins with differing results. An influence of passenger protein structure on the efficacy of transport has been noted for both OE23 and OE17. Specifically, C-terminal truncations of pOE23 impaired its transport [73] and insertion of proline residues into the fourth helical region of OE17 severely inhibited its transport (Braun and Theg, unpublished data). On the other hand, Hynds *et al*. [46] reported efficient transport of a C-terminally truncated pOE23-DHFR as well as pOE23-DHFR that was translated with amino acid analogues to destabilize its conformation. Experiments (K. Cline, in preparation) have examined the ability of Tat to transport tandem repeats of the ($Gly_4Ser$) peptide, which has been used as an unstructured linker in producing single chain antibodies [74]. Efficient transport was obtained for a chimeric protein consisting of a targeting peptide fused to 15 such repeats (75 amino acids), which theoretically could be 25 nm in length. Transport of unfolded or unstructured proteins differs from the situation with the *E. coli* Tat system, which rejects unfolded substrates [67].

### *2. Biochemical Requirements*

#### a. Soluble Factors and NTPs Are not Required

Thylakoid Tat protein transport of purified precursors into washed thylakoids [16, 69, 75, 76] is as efficient in the absence as in the presence of stromal extract [16]. It is possible or even likely that after import into the chloroplast, Tat pathway precursors interact with stromal chaperones, for example, to assist folding, but these appear unnecessary for protein translocation per se. Tat transport also occurs efficiently without NTPs. Careful removal of all NTPs from the assay mixture and inclusion of nonhydrolyzable NTP analogues does not impair transport [12]. Furthermore, the measured ATP concentration in the lumen is less than 1 μM [12], and the identified Tat components lack nucleotide-binding motifs. In this respect, Tat systems differ from virtually all other protein translocation systems.

#### b. Tat Transport Requires the Thylakoidal $\Delta\mu^{H^+}$

Tat transport *in vitro* requires energy in the form of the thylakoidal $\Delta\mu^{H^+}$. The $\Delta\mu^{H^+}$ is typically generated with light through photosynthetic electron transport, but can also be generated by reverse action of the ATP synthase [12]. With isolated thylakoids, the $\Delta\mu^{H^+}$ is usually parsed primarily into the

pH gradient [11, 12]. Thus, dissipating the ΔpH with selective ionophores is generally sufficient to prevent transport.

A careful analysis of the energetics of the Tat system arrived at three major conclusions [76]. First, there is a threshold ΔpH for transport that varies with the substrate protein and beyond which the rate of protein transport is linearly dependent on the driving force. This is the expected behavior of a chemiosmotic system, in which the threshold represents the minimum thermodynamic potential below which there is not enough energy in the gradient to perform the measured work, in this case, protein transport. Second, a proton counterflow occurs during protein translocation. Again, this is a requirement of chemiosmotic coupling, as the energetic content of the gradient must be traded for the work of protein translocation. What was unexpected was the magnitude of the proton counterflow, some 80,000 $H^+$ per protein transported. Finally, it was observed that the number of protons drained from the gradient per protein translocated is constant at varying ΔpHs, indicating that the proton flow is not due to a leak, the rate of which would certainly vary with the concentration gradient of protons. Indeed, it was calculated that over 90% of the proton counterflow is mechanistically linked to protein translocation. As will be seen in section V.C.2.b, the ΔpH is also required for assembly of the translocase, although it is not known if proton counterflow accompanies this process.

A challenge to the *in vivo* requirement for the ΔpH has come from experiments with the alga *C. reinhardtii* [77]. Using a variety of experimental approaches, including a mutant incapable of generating a pH gradient, the authors showed that a ΔpH per se is not necessary for Tat protein transport. However, they had previously reported that a considerable Δψ is generated in the mutant alga and could not rule out the possibility that Δψ can be used to power transport in place of the ΔpH [78]. Experiments with isolated thylakoids, in which the ΔpH was deliberately reduced to require that more of the $\Delta\mu^{H^+}$ be carried by the electric potential, show that the Δψ can contribute to the driving force for transport (Braun, Davis, and Theg, submitted for publication).

A second *in vivo* study employed tobacco protoplasts and overexpressed precursor proteins and concluded that neither the Δψ nor the ΔpH are required for Tat transport [79]. Here, nigericin and valinomycin were applied to the protoplasts to dissipate the ΔpH and Δψ, respectively, and Tat transport was observed under conditions where no thylakoidal potentials were expected. Although carefully planned and executed, these experiments are subject to some uncertainties. First, the Δψ was not measured and the ΔpH was assessed by loss of the rapid phase of the nonphotochemical fluorescence quenching (NPQ) signal. As with other methods that measure the ΔpH, NPQ is not sensitive to the pH gradient when it drops

to low but significant values [80]. In this regard, experiments (Braun and Theg, in preparation) have demonstrated ionophore-sensitive Tat transport in isolated thylakoids long after cessation of illumination and when a pH gradient could not be detected spectroscopically. Also, the time frame during which protein transport was measured in the tobacco protoplast experiments was in the range of one to several hours. This not only allows very inefficient transport reactions to be scored as successful but also extends the time past which the applied ionophores are effective. For instance, it has been shown that in some circumstances the membrane electric potential can recover in isolated thylakoids in the presence of valinomycin under steady state illumination [81]. These considerations emphasize the complexities of thylakoid energetics and caution that it is premature to conclude that Tat pathway energetics are substantially different *in vivo* than *in vitro*. Clearly, this is an area requiring more experimentation before it will be settled.

If, as it seems, the Tat system can use the full $\Delta\mu^{H^+}$ to power protein transport, then there are important mechanistic implications. Mitchell [82] pointed out that, although the $\Delta$pH and $\Delta\psi$ provide thermodynamically equivalent forces, there is no obvious mechanistic reason why the two should lead to equivalent rates of a given chemiosmotic process. Mitchell proposed the existence of a ''proton well'' as a device that would enable a process to employ either the $\Delta$pH or $\Delta\psi$. A proton well would convert the electric potential into a chemical potential, that is the $\Delta$pH, by allowing free passage of protons much of the way across the membrane. Such a proton well is observed in the structure of the F0/F1 $H^+$-ATPases of energy-transducing membranes, wherein the F0 portion of the enzyme is a proton channel. This consideration suggests that a similar proton channel will be found within the structure of the Tat machinery to serve as an energy-transducing proton well.

### c. Components of the Tat Machinery: Three Membrane Proteins Are Required for Transport

The above-mentioned capabilities and requirements describe a system that can transduce the electrochemical energy of the $\Delta\mu^{H^+}$ into the mechanical energy necessary to transport folded proteins of varying diameter or completely unfolded proteins across a sealed membrane [75]. These tasks appear to be accomplished by only three membrane proteins (Figure 18.3). The first identified Tat component, maize Hcf106 [49, 62] came from a genetic screen for plants selectively defective in transport of proteins

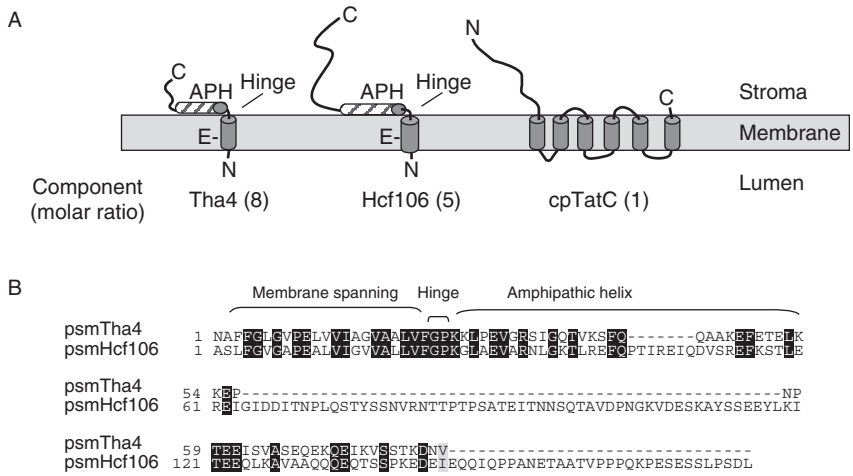



FIG. 18.3. Components of the translocation machinery for the Tat pathway in chloroplasts. (A) Diagram of the predicted structural features of the thylakoid Tat pathway the components. The molar ratios of components in pea thylakoid membranes are in parentheses. Hcf106 and Tha4 both possess a conserved transmembrane glutamate (E), a predicted flexible hinge, and predicted amphipathic helix (APH). (B) Amino acid sequence alignment of the mature form of pea Tha4 (psmTha4) with the mature form of pea Hcf106 (psmHcf106). Alignment was made with ClustalW and the shading by Boxshade.

known to require only the ΔpH. This was quickly followed by identification of *E. coli* Tat components [83–85], which are encoded in the same operon, and the remaining thylakoid components Tha4 [86] and cpTatC [87, 88]. The bacterial orthologues of Tha4, Hcf106, and cpTatC are called TatA, TatB, and TatC, respectively. In pea thylakoid membranes, Tat components have been estimated at 140,000 Tha4, 95,000 Hcf106, and 18,000 cpTatC per chloroplast equivalent or an 8:5:1 ratio [87]. A similar, but even more pronounced ratio has been determined for *E. coli* Tat components [72].

Hcf106 and Tha4 are similar structurally and on the amino acid sequence level, yet they seem to play different roles in the transport process (Figure 18.3). Both proteins are anchored to the membrane by an amino proximal transmembrane domain, followed by predicted glycine–proline hinge, a stroma exposed amphipathic helix, and an acidic C-terminal tail [58]. cpTatC is an integral membrane protein with six predicted membrane spanning helices and its N- and C-termini exposed to the stroma [87]. All three components are required for protein transport *in vitro* as assessed by

antibody inhibition studies. Pretreatment of pea thylakoids with antibodies to any single component specifically disables the Tat pathway [44, 87]. Antibodies to Hcf106 or cpTatC inhibit precursor binding and protein translocation; antibodies to Tha4 inhibit translocation but are without effect on binding [29].

Several arguments have been advanced for the notion that there are no other core components of the translocase. These include the observation that although genes for TatA, TatB, and TatC are frequently genetically linked, no uncharacterized genes are consistently linked to the Tat genes in sequenced prokaryotic genomes [65]. Second, purified *E. coli* Tat complexes contains only TatA, TatB, and TatC [89, 90]. Nevertheless, until transport is reconstituted with proteoliposomes-containing purified components, this will remain an open question. In addition to the essential components of the Tat machinery, other factors may regulate or stimulate Tat transport. For example, a genetic selection in *E. coli* identified PspA as a protein that relieved the pressure of overexpressing a Tat substrate [91]. Studies with the thylakoid system have shown that the cyanobacterial and chloroplast orthologues of PspA, called Vipp1, stimulate Tat transport *in vitro* to ~150% of control (Theg, unpublished data).

Structure–function studies of thylakoid Tat components have been limited to analyses with *in vitro* assays. An *in vitro* substitution assay was developed to replace endogenous Tha4 with recombinant *in vitro* translated Tha4 [92]. It involves anti-Tha4 antibody treatment to inactivate the endogenous protein, followed by *in vitro* integration to introduce the recombinant protein. Mutational analysis with this assay showed that the Tha4 amphipathic helix and transmembrane domain, but not the carboxyl tail, are essential for function [92]. This result is similar to findings for *E. coli* TatA [93]. Replacement of the Tha4 transmembrane domain with the Hcf106 transmembrane domain resulted in a partially functional Tha4 [92]. Cysteine-scanning mutagenesis of the Tha4 transmembrane and hinge region had minor effects on function with a single exception. Substitution of transmembrane glutamate E10 totally eliminated function [94]. Several other amino acid substitutions of Tha4's E10 were also nonfunctional; even substitution with aspartate was largely ineffective [92]. The Tha4 E10 residue was shown to be essential for Tha4 assembly into the translocase (see section V.C.2.*b*), but certainly is also a candidate for sensing or transducing the ΔpH. Hcf106 contains a conserved transmembrane glutamate in the same relative position as the Tha4 E10 (Figure 18.3). *In vitro* integrated Hcf106 assembles with cpTatC [58] and, interestingly, substitution of the Hcf106 transmembrane glutamate with glutamine prevented this assembly [58]. This implies that the Hcf106 transmembrane glutamate is also essential for function.

## C. Operation of the Tat System

The thylakoid Tat transport process has been staged into two steps; precursor protein binding to the membrane and protein translocation. Precursor binding is generally assayed in the absence of the ΔpH. Translocation is then initiated by energizing the membrane.

### *1. Precursor Protein Binding*

Precursors show differing ability to bind to thylakoids and Tat components that varies with the specific precursor and the method of preparation. Some precursors bind in a highly reversible manner. For example, association of the OE23 precursor with Tat components can only be detected if chemical cross-linkers are used to stabilize the interaction [95]. On the other hand, several laboratories have shown that certain precursors bind irreversibly to thylakoids [27, 69, 96]. Two studies have demonstrated precursor binding to the lipid bilayer. Others have demonstrated a primary association with Tat components.

#### a. Precursor Protein Binding to the Lipid Bilayer

Bacterially expressed and purified pOE17 (Figure 18.2) bound to thylakoids or liposomes in a salt or alkaline-resistant interaction [69]. At least some of this binding was productive in the sense that energizing the membranes led to transport of the bound precursor. Similarly, an *in vitro* translated chimeric precursor called 16/23 bound to protease-treated thylakoids or to liposomes [96]. Proteolysis of the bound 16/23 precursor produced a 14-kDa fragment. This latter result was interpreted to mean that the first stage of precursor association with thylakoids involves insertion of the signal peptide and some of the mature protein into the lipid bilayer.

#### b. Precursor Protein Binding to a cpTatC–Hcf106 Receptor Complex

Other precursors bind primarily to Tat components. In thylakoids that are not transporting proteins, Tat components exist in several different complexes. cpTatC and Hcf106 comprise a ~700-kDa complex as determined by BN-PAGE analysis and coimmunoprecipitation with digitonin-solubilized thylakoids [29]. Cross-linking of thylakoids with disuccinimidyl suberate resulted in a ladder of cpTatC-containing bands up to at least 600 kDa on an sodium dodecyl sulfate (SDS) denaturing gel (K. Cline, unpublished data). Because the ladder of bands can be immunoprecipitated with anti-Hcf106 IgGs, this confirms that the large cpTatC–Hcf106 complex is the native complex rather than an artifact of detergent solubilization. Partial purification of the thylakoid complex (H. Mori and K. Cline, unpublished data) and analogy with the large *E.coli* Tat complex [90, 97] suggests

that the 700-kDa complex contains multiple copies of cpTatC and Hcf106 in a 1:1 ratio. Tha4 exists in a separate complex that appears to be a Tha4 homo-oligomer (C. Dabney-Smith and K. Cline, in preparation). There is also a separate pool of Hcf106 with unknown composition or function. Chemical cross-linking studies have confirmed that this organization exists *in situ* [29, 95].

Precursors that bind irreversibly to Tat components include several recombinant proteins [27, 29, 98], including truncated precursor to OE17 (tOE17), a truncated form of the OE17 precursor (Figure 18.2). These precursors bind to the 700-kDa cpTatC–Hcf106 complex (termed the receptor complex) [27, 29, 95, 98] and can be cross-linked to cpTatC and Hcf106, but not Tha4 [29, 95, 98]. Precursors that bind to the cpTatC–Hcf106 complex are >85% productive [27, 98] and their binding is strictly dependent on the twin arginine motif and an uninterrupted hydrophobic core region of the signal peptide [29, 95, 98]. The stoichiometry of precursor to receptor complex components under saturating conditions is yet to be determined, but the cpTatC–Hcf106 complex is presumably multivalent. In this regard, a kinetic study of Tat transport of the OE17 precursor found evidence for cooperativity with respect to substrate concentration, which is consistent with a multivalent cpTatC–Hcf106 complex [30].

Fine mapping of tOE17-component interactions by site directed photocross-linking showed that signal peptide residues proximal to the RR strongly interact with cpTatC, whereas residues in the hydrophobic core interact with Hcf106 [98]. Cross-linking products were not obtained when the photoreactive group was placed in the early mature domain of tOE17, nor was any cross-linking to Tha4 observed. These results are similar to the pattern of interactions determined for preSufI and *E. coli* Tat components [99] with an interesting difference. In the *E. coli* study, cross-linking to TatA was also observed when the membranes were energized.

An unexpected result of the photocross-linking analysis was that substitution of some tOE17 signal peptide residues with the photoreactive (Tmd) Phe, an analogue of phenylalanine, altered the nature of the interaction between tOE17 and the cpTatC–Hcf106 complex. Whereas tOE17 binds to cpTatC–Hcf106 in a salt-sensitive interaction that is somewhat unstable during BN-PAGE, the Tmd-Phe substituted tOE17 bound in a salt-resistant interaction that was completely stable to BN-PAGE [98]. A similarly tight binding precursor was obtained by a single phenylalanine substitution of the RR proximal valine-20 of tOE17 (Figure 18.2, F. Gerard and K. Cline, in preparation). tOE17F-20 also appeared to bind more deeply into the cpTatC–Hcf106 complex. The accessibility of the thylakoid-bound precursor signal peptide regions was determined by engineering factor Xa cleavage sites into the precursor. Whereas both the amino- and carboxyl regions

flanking the signal peptide were accessible to factor Xa protease with thylakoid-bound tOE17, the same regions including up to 40 residues into the mature OE17 were inaccessible to factor Xa with thylakoid-bound tOE17F-20 (F. Gerard and K. Cline, in preparation).

So what is the physiological significance of this range of binding interactions? Two observations suggest that tight binding represents an advanced stage of the transport reaction. First, tight binding precursors are the most efficient transport substrates [27]. Second, thylakoid-bound tOE17 was more salt resistant and less exposed to factor Xa when thylakoids were energized with the ΔpH (F. Gerard and K. Cline, in preparation). A caveat to this type of experiment is that it could only be conducted with thylakoids pretreated with anti-Tha4 IgGs to prevent translocation [29, 92]. One possible interpretation of these results is that strong binding of cpTatC–Hcf106 to the signal peptide enables it to move the precursor across the membrane. In fact, a study demonstrated efficient transport of tOE17 despite its covalent attachment to cpTatC via an RR-proximal residue of the signal peptide [98].

### *2. The Translocation Step*

#### a. Simultaneous Translocation of the Entire Protein Through Some Sort of Permeation Pathway

In contrast to the precursor-binding step, which is stable, the translocation step is transient and difficult to characterize. Nevertheless, several observations give some insight into the process. First, the entire mature domain is transported at once. Hashimoto *et al.* [48] showed that during transport of the natural precursor protein preOE23, cleavage of the signal peptide occurred simultaneously with acquisition of protection to exogenous protease. This differs from the progressive linear transport by the Sec system (above). Second, several experiments with recombinant precursors demonstrated access of the signal peptide cleavage site to the lumen, while part or all of the mature passenger protein was still accessible to proteases added from the stromal side of the membrane. One such experiment used a Tat pathway precursor bound to avidin via a biotin moiety incorporated into the C-terminus [100]. After transport, this substrate was associated with the membrane, cleaved by the thylakoid signal peptidase, but the mature domain of the substrate was accessible to proteases added from the *cis* side of the membrane. A second experiment placed the unstructured (Gly4Ser) repeat peptide (above) between pOE17 and protein A. After transport, this substrate, with its signal peptide removed, spanned the membrane with the OE17 moiety in the lumen and the protein A moiety accessible to proteases on the *cis* side of the membrane (K. Cline, in

preparation). An *in vivo* experiment in which Tat substrates were transiently overexpressed in tobacco protoplasts found processed Tat substrates in the stromal fraction of recovered chloroplasts [101]. One interpretation of this result is that the precursor was transported far enough to allow signal peptide processing but not far enough to deliver the passenger protein to the lumen, and the processed protein resolved back into the stromal space.

A third observation is that the N-terminus of the signal peptide remains on the *cis* side of the membrane following transport of the mature domain. For example, pOE17 that contained a large protein fused to its N-terminus was efficiently transported to the lumen, while the N-terminal fusion protein remained on the *cis* side of the membrane [102]. In addition, the chimeric precursor 16/23, which was transported without signal peptide cleavage, exposed its N-terminus to protease added to the *cis* side of the membrane and remained associated with the Tat complex [103].

Transport of folded proteins of varied diameter across a sealed membrane poses a substantial mechanistic challenge. Several models have been proposed including an endocytic mechanism, a gated box [104], a flexible dynamic channel [65], and direct transport through the lipid bilayer [105]. The above observations argue against an endocytic mechanism, but do not specifically address the other possibilities. However, the fact that the N-terminus of the signal peptide remains on the *cis* side of the membrane evokes a transport mechanism in which the precursor pivots on a point defined by the RR region of the signal peptide and its attachment site on cpTatC (Figure 18.4).

### b. A Tha4 Oligomer Assembles with the Precursor–cpTatC–Hcf106 Complex to Form the Putative Translocase

Tat components assemble into a supercomplex during translocation and this may reflect the mechanism of protein transport (Figure 18.4). Tha4 can be chemically cross-linked to the cpTatC–Hcf106 complex in the presence of precursor and a thylakoidal ΔpH [95]. A synthetic signal peptide was sufficient to induce this assembly, but a KK-mutant precursor did not. Time course analysis showed that Tha4 assembled with the precursor receptor complex in advance of translocation and disassembled after the precursor had been transported [95]. These results imply that signal peptide binding triggers the formation of a ''translocase'' complex that transports the precursor and then disassociates to reset the system. Because Tha4 plays no role in precursor recognition, this result points to a specific and essential role for Tha4 in the translocation step.

The chemical cross-linking approach yielded little information on the size, organization, or stoichiometry of components in the putative translocase. However, oxidative disulfide cross-linking between singly Cys-substituted

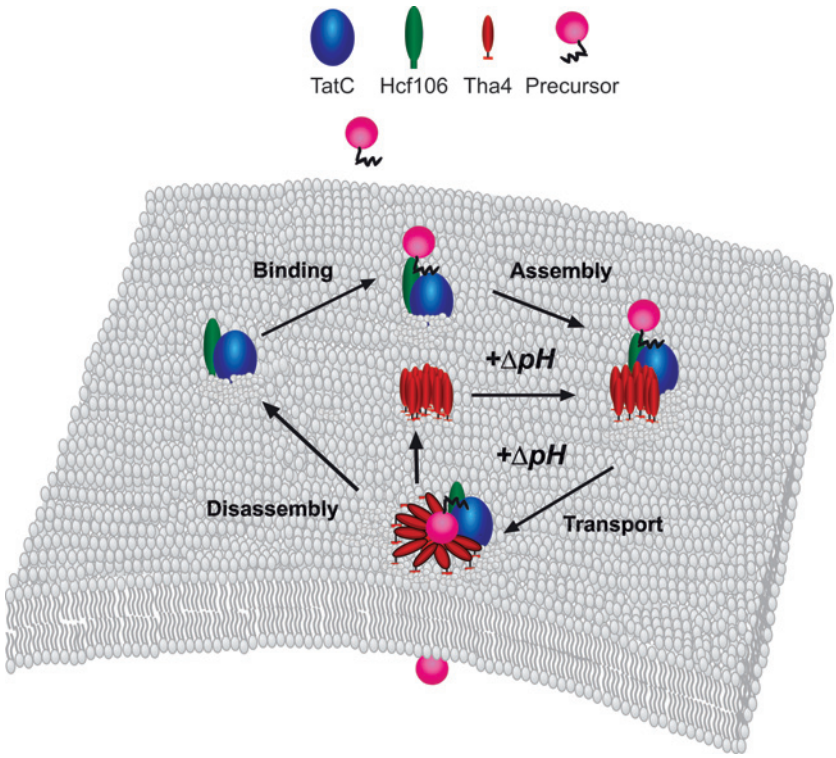



FIG. 18.4. Model for thylakoid Tat protein translocation with Tha4 providing a passive trapdoor protein-conducting channel. The figure depicts a cyclical model for the operation of the thylakoid Tat apparatus. Experimental data supports several aspects of this model: (1) cpTatC–Hcf106 and Tha4 are present as independent complexes (the additional Hcf106 complex is not shown); (2) the folded precursor protein binds to cpTatC–Hcf106, which triggers docking of a Tha4 oligomer that undergoes conformational rearrangement; and (3) the protein is translocated and the Tha4 oligomer dissociates. Speculative aspects of the model are that the Tha4 oligomer amphipathic helices fold into the lipid bilayer and provide a passive permeation pathway for the translocated protein. Other investigators have suggested that Tha4 (TatA) provides a fixed gated channel or that Tha4 (TatA) serves to "weaken" the bilayer, thereby lowering the resistance to protein movement, or that Tha4 (TatA) adjusts the size of a channel provided by the cpTatC–Hcf106 (TatBC) complex (see text). Although the cpTatC–Hcf106 complex is shown as a heterodimer in the figure, experimental evidence indicates that it contains multiple copies of cpTatC and Hcf106.

Tha4 gave some insight into Tha4 organization during transport [94]. Tha4–Tha4 interactions were found to occur at comparable sites across the entire Tha4 molecule. Interactions through the transmembrane domain occurred even with transport-inactive membranes. Interactions mediated through the amphipathic helix and carboxyl tail occurred were strictly dependent

on precursor and the thylakoidal ΔpH. Limited cross-linking with disuccinimidyl suberate, which reacts with the lysine residues in the amphipathic helix and carboxyl tail, showed that interactions through these domains result in oligomers at least as large as octamers [94]. This result has now been confirmed by disulfide cross-linking experiments with doubly Cys-substituted Tha4. Tha4 with double Cys substitutions in the transmembrane domain yielded oligomers at least as large as decamers with transport-inactive thylakoids, whereas double Cys substitutions in the stromally exposed regions yielded Tha4 oligomers of a similar size, but only under transport conditions (C. Dabney-Smith and K. Cline, in preparation). Tha4 oligomers could be induced with a synthetic signal peptide but not a KK-mutant precursor, indicating that the mature domain of the precursor does not influence Tha4–Tha4 interaction or the apparent size of the oligomer. Furthermore, oligomerization was prevented by pretreating thylakoids with anti-cpTatC IgGs, suggesting a requirement for Tha4 interaction with the cpTatC–Hcf106 complex [94].

One interpretation for these results is that Tha4 exists in nonactive thylakoid membranes as an oligomer that associates through its transmembrane domain. Signal peptide binding to the cpTatC–Hcf106 complex exposes a docking site for Tha4 and docking induces a conformational reorganization that brings Tha4s amphipathic helices and carboxyl tails into direct contact. This might be accomplished by alignment of the amphipathic helices in the interfacial region of the lipid bilayer. Tha4's amphipathic helix is predicted to be surface active [94].

### D. Models for the Tat Translocase and Future Directions

Several models have been proposed for Tat transport systems in bacteria and thylakoid membranes [24, 25, 65, 94, 99, 106]. Most models include binding of the precursor to cpTatC–Hcf106 (TatBC) and the subsequent recruitment of Tha4 (TatA) to form the translocase. A different order of steps has emerged from studies of the *Bacillus subtilis* Tat system, where a TatA oligomer appears to bind the precursor in the cytosol and target it to membrane anchored-TatC [107].

Two models for the translocation step suggest that the precursor is transported through a fixed channel. These models are based primarily on single particle imaging of isolated Tat complexes. Robinson and Bolhuis [25] suggest that the TatBC complex provides the basic channel and that TatA binds TatBC and adjusts the channel diameter to fit the precursor [25]. This model draws on images of the isolated *E. coli* TatBC complex, which appear to show channel-like openings [108]. A second ''channel'' model suggests that TatA oligomers provide size-appropriate channels for

the substrate being transported [89]. This model is based on single particle imaging of purified TatA that appears to show a collection of ring-like complexes with varying diameters [89]. Both channel models imply a mechanism that can sense the dimensions of the substrate and either adjust the channel opening or select the appropriate TatA complex, such that a leak-free translocation pathway is obtained.

Two significantly different models for translocation suggest that precursors do not pass through a channel in the accepted sense, but are carried across the lipid bilayer in a manner that is facilitated by Tha4 (TatA). Bruser and Sanders [105] suggest that the presence of massed TatA proteins near the precursor-bound TatBC complex induces a locally ''weakened'' bilayer through which the precursor could be pulled. Dabney-Smith *et al.* [94] propose a mechanistic explanation for such membrane ''weakening.'' Specifically they suggest that the Tha4 oligomer facilitates transport by allowing its amphipathic helices to fold into the bilayer in response to a mechanical force on the precursor. This could provide a passive and transient gate in which the hydrophilic/charged faces of the amphipathic helices would contact the folded precursor domain (Figure 18.4). Because the size and shape of the substrate would dictate the number of Tha4 protomers that undergo infolding, the chances of ion leakage would be minimized. This model is based on the finding that the size of the cpTatC–Hcf106 associated Tha4 oligomer appears independent of the size of the precursor, or even the presence of the mature domain [94]. It also draws on analogy with the pore-forming amphipathic helical peptides [109] that concentrate on the membrane surface before forming a transmembrane permeation pathway, and on observations that *E. coli* TatA undergoes topology inversion as a result of protein translocation [110]. If proteins pass through the bilayer, then one might expect that specific lipids play crucial roles in transport efficiency. In a study, Ma and Browse [111] found that thylakoids from *Arabidopsis* mutants with more highly saturated fatty acids were impaired in Tat pathway transport but enhanced in Sec pathway transport. In addition, *E. coli* Tat transport has been found to be dependent on anionic and nonbilayer forming phospholipids [112].

Only one of the models addresses the mechanical force required for transport. Bruser and Sanders [105] suggest that TatC pulls the precursor across the membrane. Presumably this is predicated on the demonstrated strong interaction between TatC and the signal peptide (above, section V. C.1.*b*) and on a hypothetical major conformational change of TatC. Precedent exists for such a mechanism; Colicin Ia undergoes a $\Delta\psi$-induced conformational change that moves a segment across the bilayer and is capable of carrying peptides with it [113]. Such a mechanism could work for folded protein domains, although there would be a limitation on the length of unfolded domain that could be

transported by a pulling mechanism unless TatC could attach at multiple points along the substrate.

At present the lack of information about the translocation step precludes a knowledge-based evaluation of the various models. However, each model makes predictions that hopefully will be tested in the near future. For example, the channel models predict that arrested substrates such as those described above (section V.C.2.*a*) would be trapped in a Tat component-lined channel, whereas the ''weakened'' membrane and the Tha4-facilitated membrane models predict that arrested substrates would be trapped in or across the bilayer. Because these are hydrophilic segments, they presumably would resolve *cis* or *trans* to one or the other side of the membrane or be stuck in a metastable state across the bilayer, depending on the arrangement of folded and unfolded domains. Testing these and other predictions of the models are important goals for future efforts, as is determining which component if any contacts the mature domain of the precursor during the translocation step. Finally, determining the identity of the proton carrier during transport and the manner by which this is coupled to translocation will be a crucial step in unraveling the mechanism of Tat transport.

## Acknowledgments

The authors gratefully acknowledge support by the National Institutes of Health (K.C.) and the Department of Energy (S.M.T.) during the preparation of this chapter. The authors thank Carole Dabney-Smith, Fabien Gerard, and Nicolai Braun for critical review of the chapter.

## References

1. Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet* 19:47–56.
2. Peltier, J.B., Emanuelsson, O., Kalume, D.E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D.A., Soderberg, L., Roepstorff, P., von Heijne, G., and van Wijk, K.J. (2002). Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. *Plant Cell* 14:211–236.
3. Schubert, M., Petersson, U.A., Haas, B.J., Funk, C., Schroder, W.P., and Kieselbach, T. (2002). Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. *J Biol Chem* 277:8354–8365.
4. Sugiura, M. (1992). The chloroplast genome. *Plant Mol Biol* 19:149–168.
5. Race, H.L., Herrmann, R.G., and Martin, W. (1999). Why have organelles retained genomes? *Trends Genet* 15:364–370.
6. Keegstra, K., and Cline, K. (1999). Protein import and routing systems of chloroplasts. *Plant Cell* 11:557–570.
7. Jarvis, P., and Robinson, C. (2004). Mechanisms of protein import and routing in chloroplasts. *Curr Biol* 14:R1064–R1077.

8. Smith, M.D., Fitzpatrick, L., Keegstra, K., and Schnell, D.J. (2002). "*In vitro* analysis of chloroplast protein import." Current Protocols in Cell Biology (Vol. 11.16), Hoboken: John Wiley and Sons, Inc.
9. Cline, K., Henry, R., Li, C.-J, and Yuan, J. (1992). Pathways and intermediates for the biogenesis of nuclear-encoded thylakoid proteins, in N. Murata, (ed.), *Research in Photosymthesis* (Vol. III), pp. 149–156, Amsterdam: Kluwer Academic Publishers.
10. Reed, J.E., Cline, K., Stephens, L.C., Bacot, K.O., and Viitanen, P.V. (1990). Early events in the import/assembly pathway of an integral thylakoid protein. *Eur J Biochem* 194:33–42.
11. Mould, R.M., and Robinson, C. (1991). A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. *J Biol Chem* 266:12189–12193.
12. Cline, K., Ettinger, W.F., and Theg, S.M. (1992). Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J Biol Chem* 267:2688–2696.
13. Flugge, U.I., and Hinz, G. (1986). Energy dependence of protein translocation into chloroplasts. *Eur J Biochem* 160:563–570.
14. Henry, R., Kapazoglou, A., McCaffery, M., and Cline, K. (1994). Differences between lumen targeting domains of chloroplast transit peptides determine pathway specificity for thylakoid transport. *J Biol Chem* 269:10189–10192.
15. Knott, T.G., and Robinson, C. (1994). The secA inhibitor, azide, reversibly blocks the translocation of a subset of proteins across the chloroplast thylakoid membrane. *J Biol Chem* 269:7843–7846.
16. Cline, K., Henry, R., Li, C., and Yuan, J. (1993). Multiple pathways for protein transport into or across the thylakoid membrane. *EMBO J* 12:4105–4114.
17. Hulford, A., Hazell, L., Mould, R.M., and Robinson, C. (1994). Two distinct mechanisms for the translocation of proteins across the thylakoid membrane, one requiring the presence of a stromal protein factor and nucleotide triphosphates. *J Biol Chem* 269:3251–3256.
18. Yuan, J., Henry, R., McCaffery, M., and Cline, K. (1994). SecA homolog in protein transport within chloroplasts: evidence for endosymbiont-derived sorting. *Science* 266:796–798.
19. Nakai, M., Goto, A., Nohara, T., Sugita, D., and Endo, T. (1994). Identification of the SecA protein homolog in pea chloroplasts and its possible involvement in thylakoidal protein transport. *J Biol Chem* 269:31338–31341.
20. Cline, K. (2003). Biogenesis of green plant thylakoid membranes, in B. Green, and W. Panson (eds.), *Light-Harvesting Antennas in Photosynthesis. Advances in Photosynthesis and Respiration,* (Vol. 13), pp. 353–372, Dordrecht: Kluwer Academic Publishers.
21. Molik, S., Karnauchov, I., Weidlich, C., Herrmann, R.G., and Klosgen, R.B. (2001). The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts: missing link in the evolution of protein transport pathways in chloroplasts? *J Biol Chem* 276:42761–42766.
22. Summer, E.J., Mori, H., Settles, A.M., and Cline, K. (2000). The thylakoid delta pH-dependent pathway machinery facilitates RR-independent N-tail protein integration. *J Biol Chem* 275:23483–23490.
23. Cline, K., and Henry, R. (1996). Import and routing of nucleus-encoded chloroplast proteins. *Annu Rev Cell Dev Biol* 12:1–26.
24. Mori, H., and Cline, K. (2001). Post-translational protein translocation into thylakoids by the Sec and DeltapH-dependent pathways. *Biochim Biophys Acta* 1541:80–90.
25. Robinson, C., and Bolhuis, A. (2004). Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim Biophys Acta* 1694:135–147.

26. Henry, R., Carrigan, M., McCaffrey, M., Ma, X., and Cline, K. (1997). Targeting determinants and proposed evolutionary basis for the Sec and the Delta pH protein transport systems in chloroplast thylakoid membranes. *J Cell Biol* 136:823–832.
27. Ma, X., and Cline, K. (2000). Precursors bind to specific sites on thylakoid membranes prior to transport on the delta pH protein translocation system. *J Biol Chem* 275:10016–10022.
28. Chaddock, A.M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R.G., Klosgen, R.B., and Robinson, C. (1995). A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. *EMBO J* 14:2715–2722.
29. Cline, K., and Mori, H. (2001). Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J Cell Biol* 154:719–729.
30. Alder, N.N., and Theg, S.M. (2003). Protein transport via the cpTat pathway displays cooperativity and is stimulated by transport-incompetent substrate. *FEBS Lett* 540:96–100.
31. Bogsch, E., Brink, S., and Robinson, C. (1997). Pathway specificity for a delta pH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. *EMBO J* 16:3851–3859.
32. Clausmeyer, S., Klosgen, R.B., and Herrmann, R.G. (1993). Protein import into chloroplasts. The hydrophilic lumenal proteins exhibit unexpected import and sorting specificities in spite of structurally conserved transit peptides. *J Biol Chem* 268:13869–13876.
33. Osborne, A.R., Rapoport, T.A., and van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. *Annu Rev Cell Dev Biol* 21:529–550.
34. Bolhuis, A. (2004). The archaeal Sec-dependent protein translocation pathway. *Philos Trans R Soc Lond B Biol Sci* 359:919–927.
35. Veenendaal, A.K., van der Does, C., and Driessen, A.J. (2004). The protein-conducting channel SecYEG. *Biochim Biophys Acta* 1694:81–95.
36. Dalbey, R.E., and Chen, M. (2004). Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta* 1694:37–53.
37. Vrontou, E., and Economou, A. (2004). Structure and function of SecA, the preprotein translocase nanomotor. *Biochim Biophys Acta* 1694:67–80.
38. Asakura, Y., Hirohashi, T., Kikuchi, S., Belcher, S., Osborne, E., Yano, S., Terashima, I., Barkan, A., and Nakai, M. (2004). Maize mutants lacking chloroplast FtsY exhibit pleiotropic defects in the biogenesis of thylakoid membranes. *Plant Cell* 16:201–214.
39. Randall, L.L., and Hardy, S.J. (2002). SecB, one small chaperone in the complex milieu of the cell. *Cell Mol Life Sci* 59:1617–1623.
40. Kuhn, A., Stuart, R., Henry, R., and Dalbey, R.E. (2003). The Alb3/Oxa1/YidC protein family: membrane-localized chaperones facilitating membrane protein insertion? *Trends Cell Biol* 13:510–516.
41. Gerdes, L., Bals, T., Klostermann, E., Karl, M., Philippar, K., Hunken, M., Soll, J., and Schunemann, D. (2006). A second thylakoid membrane-localized Alb3/OxaI/YidC homologue is involved in proper chloroplast biogenesis in *Arabidopsis thaliana*. *J Biol Chem* 281:16632–16642.
42. Haward, S.R., Napier, J.A., and Gray, J.C. (1997). Chloroplast SecA functions as a membrane-associated component of the Sec-like protein translocase of pea chloroplasts. *Eur J Biochem* 248:724–730.
43. Ma, X. (1999). Protein targeting and translocation on the cpSec and Delta pH pathways in chloroplast thylakoids. Ph.D dissertation, pp. 1–94. Gainesville: University of Florida.

44. Mori, H., Summer, E.J., Ma, X., and Cline, K. (1999). Component specificity for the thylakoidal Sec and Delta pH-dependent protein transport pathways. *J Cell Biol* 146:45–56.
45. Yuan, J., and Cline, K. (1994). Plastocyanin and the 33-kDa subunit of the oxygen-evolving complex are transported into thylakoids with similar requirements as predicted from pathway specificity. *J Biol Chem* 269:18463–18467.
46. Hynds, P.J., Robinson, D., and Robinson, C. (1998). The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J Biol Chem* 273:34868–34874.
47. Endo, T., Kawakami, M., Goto, A., America, T., Weisbeek, P., and Nakai, M. (1994). Chloroplast protein import. Chloroplast envelopes and thylakoids have different abilities to unfold proteins. *Eur J Biochem* 225:403–409.
48. Hashimoto, A., Ettinger, W.F., Yamamoto, Y., and Theg, S.M. (1997). Assembly of newly imported oxygen-evolving complex subunits in isolated chloroplasts: sites of assembly and mechanism of binding. *Plant Cell* 9:441–452.
49. Voelker, R., and Barkan, A. (1995). Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J* 14:3905–3914.
50. Mould, R.M., Kapazoglou, A., and Gray, J.C. (2001). Assembly of cytochrome f into the cytochrome bf complex in isolated pea chloroplasts. *Eur J Biochem* 268:792–799.
51. Rohl, T., and van Wijk, K.J. (2001). *In vitro* reconstitution of insertion and processing of cytochrome f in a homologous chloroplast translation system. *J Biol Chem* 276:35465–35472.
52. High, S., Henry, R., Mould, R.M., Valent, Q., Meacock, S., Cline, K., Gray, J.C., and Luirink, J. (1997). Chloroplast SRP54 interacts with a specific subset of thylakoid precursor proteins. *J Biol Chem* 272:11622–11628.
53. Zhang, L., Paakkarinen, V., Suorsa, M., and Aro, E.M. (2001). A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. *J Biol Chem* 276:37809–37814.
54. Nilsson, R., Brunner, J., Hoffman, N.E., and van Wijk, K.J. (1999). Interactions of ribosome nascent chain complexes of the chloroplast-encoded D1 thylakoid membrane protein with cpSRP54. *EMBO J* 18:733–742.
55. Ossenbuhl, F., Gohre, V., Meurer, J., Krieger-Liszkay, A., Rochaix, J.D., and Eichacker, L.A. (2004). Efficient assembly of photosystem II in *Chlamydomonas reinhardtii* requires Alb3.1p, a homolog of Arabidopsis ALBINO3. *Plant Cell* 16:1790–1800.
56. Klostermann, E., Droste Gen Helling, I., Carde, J.P., and Schunemann, D. (2002). The thylakoid membrane protein ALB3 associates with the cpSecY-translocase in *Arabidopsis thaliana*. *Biochem J* 368:777–781.
57. Roy, L.M., and Barkan, A. (1998). A SecY homologue is required for the elaboration of the chloroplast thylakoid membrane and for normal chloroplast gene expression. *J Cell Biol* 141:385–395.
58. Fincher, V., Dabney-Smith, C., and Cline, K. (2003). Functional assembly of thylakoid deltapH-dependent/Tat protein transport pathway components *in vitro*. *Eur J Biochem* 270:4930–4941.
59. Inoue, K., Baldwin, A.J., Shipman, R.L., Matsui, K., Theg, S.M., and Ohme-Takagi, M. (2005). Complete maturation of the plastid protein translocation channel requires a type I signal peptidase. *J Cell Biol* 171:425–430.
60. Steiner, J.M., and Loffelhardt, W. (2005). Protein translocation into and within cyanelles (review). *Mol Membr Biol* 22:123–132.
61. Tour, O., Meijer, R.M., Zacharias, D.A., Adams, S.R., and Tsien, R.Y. (2003). Genetically targeted chromophore-assisted light inactivation. *Nat Biotechnol* 21:1505–1508.
62. Settles, A.M., Yonetani, A., Baron, A., Bush, D.R., Cline, K., and Martienssen, R. (1997). Sec-independent protein translocation by the maize Hcf106 protein. *Science* 278:1467–1470.

63. Berks, B.C. (1996). A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* 22:393–404.
64. Dilks, K., Rose, R.W., Hartmann, E., and Pohlschroder, M. (2003). Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* 185:1478–1483.
65. Berks, B.C., Sargent, F., and Palmer, T. (2000). The Tat protein export pathway. *Mol Microbiol* 35:260–274.
66. Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L.F. (1999). Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem* 274:13223–13228.
67. DeLisa, M.P., Tullman, D., and Georgiou, G. (2003). Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci USA* 100:6115–6120.
68. Creighton, A.M., Hulford, A., Mant, A., Robinson, D., and Robinson, C. (1995). A monomeric, tightly folded stromal intermediate on the delta pH-dependent thylakoidal protein transport pathway. *J Biol Chem* 270:1663–1669.
69. Musser, S.M., and Theg, S.M. (2000). Characterization of the early steps of OE17 precursor transport by the thylakoid DeltapH/Tat machinery. *Eur J Biochem* 267:2588–2598.
70. Clark, S.A., and Theg, S.M. (1997). A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol Biol Cell* 8:923–934.
71. Marques, J.P., Dudeck, I., and Klosgen, R.B. (2003). Targeting of EGFP chimeras within chloroplasts. *Mol Genet Genomics* 269:381–387.
72. Berks, B.C., Palmer, T., and Sargent, F. (2003). The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* 47:187–254.
73. Roffey, R.A., and Theg, S.M. (1996). Analysis of the import of carboxyl-terminal truncations of the 23-kilodalton subunit of the oxygen-evolving complex suggests that its structure is an important determinant for thylakoid transport. *Plant Physiol* 111:1329–1338.
74. Kortt, A.A., Malby, R.L., Caldwell, J.B., Gruen, L.C., Ivancic, N., Lawrence, M.C., Howlett, G.J., Webster, R.G., Hudson, P.J., and Colman, P.M. (1994). Recombinant anti-sialidase single-chain variable fragment antibody. Characterization, formation of dimer and higher-molecular-mass multimers and the solution of the crystal structure of the single-chain variable fragment/sialidase complex. *Eur J Biochem* 221:151–157.
75. Teter, S.A., and Theg, S.M. (1998). Energy-transducing thylakoid membranes remain highly impermeable to ions during protein translocation. *Proc Natl Acad Sci USA* 95:1590–1594.
76. Alder, N.N., and Theg, S.M. (2003). Energetics of protein transport across biological membranes: a study of the thylakoid DeltapH-dependent/cpTat pathway. *Cell* 112:231–242.
77. Finazzi, G., Chasen, C., Wollman, F.A., and de Vitry, C. (2003). Thylakoid targeting of Tat passenger proteins shows no delta pH dependence *in vivo*. *EMBO J* 22:807–815.
78. Theg, S.M., Cline, K., Finazzi, G., and Wollman, F.A. (2005). The energetics of the chloroplast Tat protein transport pathway revisited. *Trends Plant Sci* 10:153–154.
79. Di Cola, A., Bailey, S., and Robinson, C. (2005). The thylakoid delta pH/delta psi are not required for the initial stages of Tat-dependent protein transport in tobacco protoplasts. *J Biol Chem* 280:41165–41170.
80. Schonknecht, G., Neimanis, S., Katona, E., Gerst, U., and Heber, U. (1995). Relationship between photosynthetic electron transport and pH gradient across the thylakoid membrane in intact leaves. *Proc Natl Acad Sci USA* 92:12185–12189.
81. Shahak, Y. (1982). Activation and deactivation of $H^+$-atpase in intact chloroplasts. *Plant Physiol* 70:87–91.
82. Mitchell, P. (1968). *Chemiosmotic coupling and energy transduction*, Bodmin: Glynn Research Ltd.

83. Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C., and Palmer, T. (1998). Overlapping functions of components of a bacterial Sec-independent protein export pathway. *EMBO J* 17:3640–3650.
84. Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A., and Turner, R.J. (1998). A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 93:93–101.
85. Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C., and Palmer, T. (1998). An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* 273:18003–18006.
86. Walker, M.B., Roy, L.M., Coleman, E., Voelker, R., and Barkan, A. (1999). The maize tha4 gene functions in sec-independent protein transport in chloroplasts and is related to hcf106, tatA, and tatB. *J Cell Biol* 147:267–276.
87. Mori, H., Summer, E.J., and Cline, K. (2001). Chloroplast TatC plays a direct role in thylakoid (Delta)pH-dependent protein transport. *FEBS Lett* 501:65–68.
88. Motohashi, R., Nagata, N., Ito, T., Takahashi, S., Hobo, T., Yoshida, S., and Shinozaki, K. (2001). An essential role of a TatC homologue of a DeltapH-dependent protein transporter in thylakoid membrane formation during chloroplast development in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 98:10499–10504.
89. Gohlke, U., Pullan, L., McDevitt, C.A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H.R., and Berks, B.C. (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci USA* 102:10482–10486.
90. Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C.M., and Robinson, C. (2001). TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *J Biol Chem* 276:20213–20219.
91. DeLisa, M.P., Lee, P., Palmer, T., and Georgiou, G. (2004). Phage shock protein PspA of *Escherichia coli* relieves saturation of protein export via the Tat pathway. *J Bacteriol* 186:366–373.
92. Dabney-Smith, C., Mori, H., and Cline, K. (2003). Requirement of a Tha4-conserved transmembrane glutamate in thylakoid Tat translocase assembly revealed by biochemical complementation. *J Biol Chem* 278:43027–43033.
93. Lee, P.A., Buchanan, G., Stanley, N.R., Berks, B.C., and Palmer, T. (2002). Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation. *J Bacteriol* 184:5871–5879.
94. Dabney-Smith, C., Mori, H., and Cline, K. (2006). Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *J Biol Chem* 281:5476–5483.
95. Mori, H., and Cline, K. (2002). A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase. *J Cell Biol* 157:205–210.
96. Hou, B., Frielingsdorf, S., and Klosgen, R.B. (2006). Unassisted membrane insertion as the initial step in DeltapH/Tat-dependent protein transport. *J Mol Biol* 355:957–967.
97. McDevitt, C.A., Hicks, M.G., Palmer, T., and Berks, B.C. (2005). Characterisation of Tat protein transport complexes carrying inactivating mutations. *Biochem Biophys Res Commun* 329:693–698.
98. Gerard, F., and Cline, K. (2006). Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site. *J Biol Chem* 281:6130–6135.
99. Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H.G., Brunner, J., and Muller, M. (2003). Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* 12:937–946.

100. Asai, T., Shinoda, Y., Nohara, T., Yoshihisa, T., and Endo, T. (1999). Sec-dependent pathway and DeltapH-dependent pathway do not share a common translocation pore in thylakoidal protein transport. *J Biol Chem* 274:20075–20078.
101. Di Cola, A., and Robinson, C. (2005). Large-scale translocation reversal within the thylakoid Tat system *in vivo*. *J Cell Biol* 171:281–289.
102. Fincher, V., McCaffery, M., and Cline, K. (1998). Evidence for a loop mechanism of protein transport by the thylakoid Delta pH pathway. *FEBS Lett* 423:66–70.
103. Berghofer, J., and Klosgen, R.B. (1999). Two distinct translocation intermediates can be distinguished during protein transport by the TAT (Deltaph) pathway across the thylakoid membrane. *FEBS Lett* 460:328–332.
104. Teter, S.A., and Klionsky, D.J. (1999). How to get a folded protein across a membrane. *Trends Cell Biol* 9:428–431.
105. Bruser, T., and Sanders, C. (2003). An alternative model of the twin arginine translocation system. *Microbiol Res* 158:7–17.
106. Muller, M., and Klosgen, R.B. (2005). The Tat pathway in bacteria and chloroplasts (review). *Mol Membr Biol* 22:113–121.
107. Schreiber, S., Stengel, R., Westermann, M., Volkmer-Engert, R., Pop, O.I., and Muller, J.P. (2006). Affinity of TATCd for TATAd elucidates its receptor function in the *Bacillus subtilis* tat translocase system. *J Biol Chem* 281:19977–19984.
108. Oates, J., Mathers, J., Mangels, D., Kuhlbrandt, W., Robinson, C., and Model, K. (2003). Consensus structural features of purified bacterial TatABC complexes. *J Mol Biol* 330:277–286.
109. Tossi, A., Sandri, L., and Giangaspero, A. (2000). Amphipathic, alpha-helical antimicrobial peptides. *Biopolymers* 55:4–30.
110. Gouffi, K., Gerard, F., Santini, C.L., and Wu, L.F. (2004). Dual topology of the *Escherichia coli* TatA protein. *J Biol Chem* 279:11608–11615.
111. Ma, X., and Browse, J. (2006). Altered rates of protein transport in Arabidopsis mutants deficient in chloroplast membrane unsaturation. *Phytochemistry* 67:1629–1636.
112. Mikhaleva, N.I., Santini, C.L., Giordano, G., Nesmeyanova, M.A., and Wu, L.F. (1999). Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in *Escherichia coli*. *FEBS Lett* 463:331–335.
113. Kienker, P.K., Jakes, K.S., Blaustein, R.O., Miller, C., and Finkelstein, A. (2003). Sizing the protein translocation pathway of colicin Ia channels. *J Gen Physiol* 122:161–176.

# 19

# *Chloroplast SRP/FtsY and Alb3 in Protein Integration into the Thylakoid Membrane*

RALPH HENRY[a] • ROBYN L. GOFORTH[a] • DANJA SCHÜNEMANN[b]

[a]*Department of Biological Sciences*
*601 Science Engineering*
*University of Arkansas*
*Fayetteville, AR 72701, USA*

[b]*Department of Molecular Botany*
*Ruhr-University Bochum*
*44780 Bochum, Germany*

## I. Abstract

Chloroplast signal recognition particle (cpSRP) and its receptor are stunning examples of evolutionarily conserved components that serve to bind and target proteins to the thylakoid membrane. Unlike its cytosolic counterparts, which direct proteins cotranslationally to the endoplasmic reticulum and the cytoplasmic membrane in bacteria, cpSRP has a dual function. In addition to its cotranslational role in targeting chloroplast-synthesized membrane proteins, a second structurally distinct form of cpSRP functions posttranslationally to target nuclear-encoded *l*ight-*h*arvesting *c*hlorophyll-binding proteins (LHCPs) to the thylakoid following their import into the chloroplast from the cytosol. The target of the posttranslational pathway is the translocase *Alb*ino3 (Alb3), which integrates LHCPs into the thylakoid membrane. Homologues of Alb3 in mitochondria (Oxa1p and Oxa2p) and



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25019-X

bacteria (YidC) promote integration of membrane proteins into the inner mitochondrial membrane and into the cytoplasmic membrane, respectively. Studies of the Oxa1/Alb3/YidC gene family indicate that this family of proteins is functionally diverse, serving alone or in conjunction with other components to integrate some proteins while acting as a membrane chaperone for stable assembly of other membrane proteins. Based on biochemical, structural, and genetic studies, along with studies of SRP and Alb3 homologues, a more detailed model has begun to emerge for how these archaic targeting/integration components function to promote the biogenesis of photosynthetic complexes. The model reflects both conserved and unique functions that evolved to meet protein sorting requirements impacted by the endosymbiotic event that gave rise to chloroplasts from a cyanobacterial progenitor.

## II. Introduction

Chloroplast thylakoid membranes form the internal membrane system in chloroplasts that function as a quantum-, electron-, and proton-transfer machine, essential for sustaining life on earth. The energy-generating capabilities of these membranes stem from the action of four different supramolecular protein complexes, each assembled from 14 to 26 different protein subunits originating from both plastid and nuclear genomes. The shear abundance of thylakoid membranes and photosynthetic complexes that operate in these membranes requires the thylakoid to be one of the major protein export sites of the photosynthetic cell. At least four different thylakoid export pathways originate in the stroma, which is also the site of transcription and translation for plastid-encoded proteins. In that context, the stroma is much like the bacterial cytosol or mitochondrial matrix. Yet the function of thylakoid export pathways is best understood for nuclear-encoded thylakoid proteins, which are expressed in the cytoplasm as full-length precursors and imported into the stroma where they gain access to conserved thylakoid export systems.

Based on the evolutionary origin of organelles, it is not surprising that protein export from the stroma to the thylakoid resembles export to the *e*ndoplasmic *r*eticulum (ER), to the mitochondrial inner membrane, and to the bacterial cytoplasmic membrane [1]. For example, proteins that must cross the thylakoid to reach their functional location on the luminal side of the membrane are transported by homologous *t*win-*a*rginine-*t*argeting (TAT) or *sec*retion (Sec) transport systems, similar to those that translocate bacterial proteins across the cytoplasmic membrane. Integral thylakoid proteins present a uniquely different set of localization issues that stem, in part, from their propensity to form aggregates in solution. For many

proteins that must integrate into the ER membrane, the mitochondrial inner membrane, or the bacterial cytoplasmic membrane, aggregation is avoided by mechanisms that promote membrane insertion as the polypeptide is being translated. Similar cotranslational insertion mechanisms appear to support integration of chloroplast-synthesized thylakoid proteins, the majority of which are integral membrane proteins. However, the nuclear-encoded *l*ight-*h*arvesting *c*hlorophyll-binding *p*roteins (LHCPs) are often the most abundant integral membrane proteins in chloroplast thylakoids and must enter conserved protein export pathways following their import into the chloroplast.

The question of how LHCPs find their way posttranslationally from the chloroplast envelope to the thylakoid and subsequently insert into the membrane has been the subject of research for more than 20 years. Hence, this question is a central theme of this chapter owing to the finding that a *c*hloroplast *s*ignal *r*ecognition *p*article (cpSRP) and SRP receptor homologue (cpFtsY) function in LHCP routing to the thylakoid where the protein Albino3 (Alb3) is required for stable integration of LHCPs. The SRP transport pathway in chloroplasts also serves to export chloroplast-encoded proteins to the thylakoid by an overlapping cotranslational mechanism. However, details of the localization mechanism are sparse for this set of proteins owing to difficulties associated with reconstituting their localization into isolated thylakoids. Nevertheless, based on a combination of genetic, biochemical, and structural studies, as well as studies of protein export in mitochondria, bacteria, and through the ER, a more detailed model of cpSRP-mediated protein export has begun to emerge. The model reflects both conserved and unique functions of an archaic targeting/ integration mechanism that evolved to meet protein-sorting requirements associated with the endosymbiotic event that gave rise to chloroplasts from a cyanobacterial progenitor.

## III. The General Pathway for Posttranslational Targeting of LHCPs by cpSRP

Like other nuclear-encoded thylakoid proteins, LHCPs are synthesized in the cytosol as full-length precursors (Figure 19.1). An N-terminal chloroplast targeting peptide directs their transport across two envelope membranes by a mechanism that relies on the Toc and Tic translocase (see [2, 3] for review). The chloroplast-targeting peptide is processed by a stromal processing protease during or soon after import into the chloroplast, yielding mature-sized proteins (generally ranging from 20 to 30 kDa). Subsequent localization to the thylakoid leads to integration,

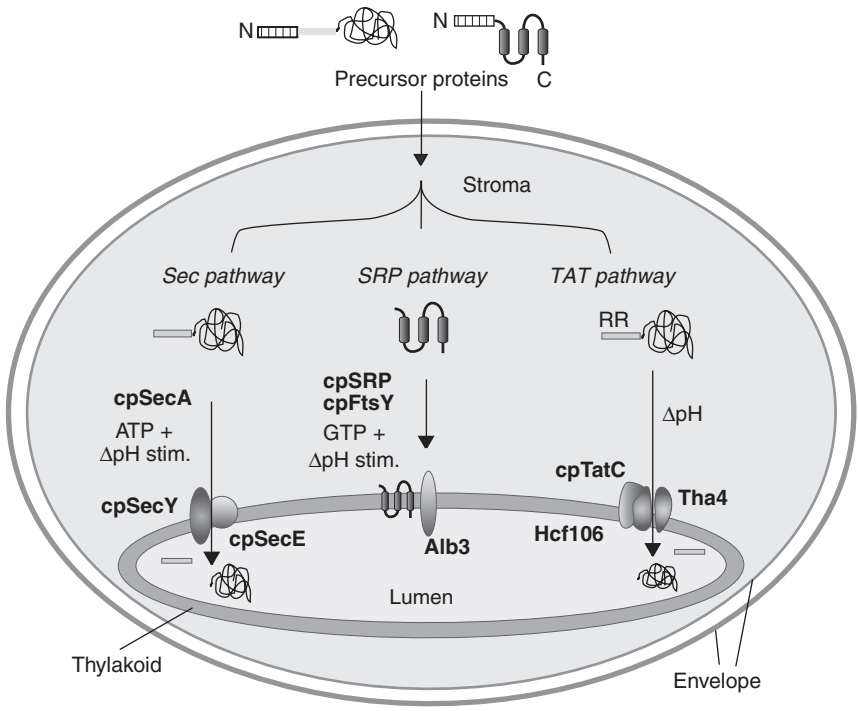



FIG. 19.1. Nuclear-encoded thylakoid proteins use evolutionarily conserved transport pathways. Precursor proteins are imported across the chloroplast envelope by the general import machinery composed of the Toc and Tic translocases (not shown, see [2, 3] for reviews). On entry into the stroma, stromal processing proteases remove the chloroplast-targeting domain. The three major pathways, Sec, SRP, and TAT, are distinguishable by their protein and energetic requirements as well as their substrate specificity. The chloroplast Sec pathway transports N-terminal signal sequence containing proteins (e.g., 33-kDa subunit of the *o*xygen-*e*volving complex, OE33) to the thylakoid lumen in a process that requires soluble cpSecA and a membrane complex of cpSecY/cpSecE [25]. The SRP pathway in chloroplasts serves to integrate the LHC family of proteins into the thylakoid membrane. The SRP pathway relies on the conserved GTPases, cpSRP54 and cpFtsY, a novel cpSRP43, and the membrane protein Alb3. TAT pathway substrate proteins (e.g., OE23) contain a twin arginine motif (RR) in the signal peptide. The membrane proteins cpTatC, Tha4, and Hcf106 as well as a ΔpH and $\Delta\psi$ *in vivo* (see [82] for review) are required for proper function of the TAT pathway. A limited set of proteins are transported by an apparently spontaneous mechanism, which has no known proteinaceous or energetic requirements [83].

chlorophyll attachment, and assembly of homotrimers, which further assemble to form peripheral *l*ight-*h*arvesting *c*omplexes (LHCs) for *p*hoto*sy*stem I (PS I) and *p*hoto*s*ystem II (PS II). The *Arabidopsis* nuclear genome codes for multiple LHC family members [4], each containing three to four membrane-spanning domains.

Early biochemical studies by Cline's group established that the pathway for thylakoid localization of LHCPs in intact chloroplasts involves the formation of a soluble pathway intermediate in the stroma termed transit complex, which has a molecular weight of ~120 kDa [5]. Transit complex formation serves to prevent LHCP aggregation and preserves LHCP in an integration competent conformation, consistent with transit complex being the substrate form of LHCP that is targeted to the thylakoid for integration. These studies, coupled with the discovery by Hoffman's group of an SRP in chloroplasts [6] and the finding that GTP is the only required nucleotide for LHCP integration into isolated thylakoids [7], set the stage for biochemical and genetic studies that have established the function of a posttranslational SRP transport pathway in chloroplasts.

## IV. Soluble and Membrane Components of the Posttranslational SRP Pathway

### A. cpSRP Is Composed of a Conserved 54-kDa GTPase and a 43-kDa Subunit Unique to Chloroplasts

Cytosolic SRPs, which function cotranslationally to target protein substrates to the ER of eukaryotes and the cytoplasmic membrane of prokaryotes, contain a 54-kDa subunit (SRP54; Ffh in *Escherichia coli*) that is central to the targeting mechanism (Figure 19.2). In addition to binding hydrophobic signal sequences of targeting substrates as they emerge from the ribosome, SRP54/Ffh possesses GTP binding and hydrolysis capabilities that are coordinated by its binding to the ribosome, its interaction with an SRP receptor at the target membrane, and its release of the targeting substrate to a protein translocase in the membrane [8–10]. Identification of a nuclear-encoded cpSRP54 homologue (cpSRP54) [6] fueled studies that showed cpSRP54 to be a component of the soluble LHCP transit complex and established its requirement for posttranslational LHCP integration into isolated thylakoids [11]. Like its pro- and eukaryotic counterparts, sequence comparison revealed the presence of a conserved NG domain that functions in GTP binding and hydrolysis along with a C-terminal M- (methionine-rich) domain, which in cytosolic SRPs interacts with the hydrophobic region of signal sequences and binds a conserved RNA moiety. It is noteworthy that the chloroplast genomes of higher plants and *Chlamydomonas* lack an indentifiable cpSRP RNA gene and analysis of cpSRP has failed to identify an RNA component, even though an SRP RNA is critical for function of cytosolic SRPs. However, the absence of an

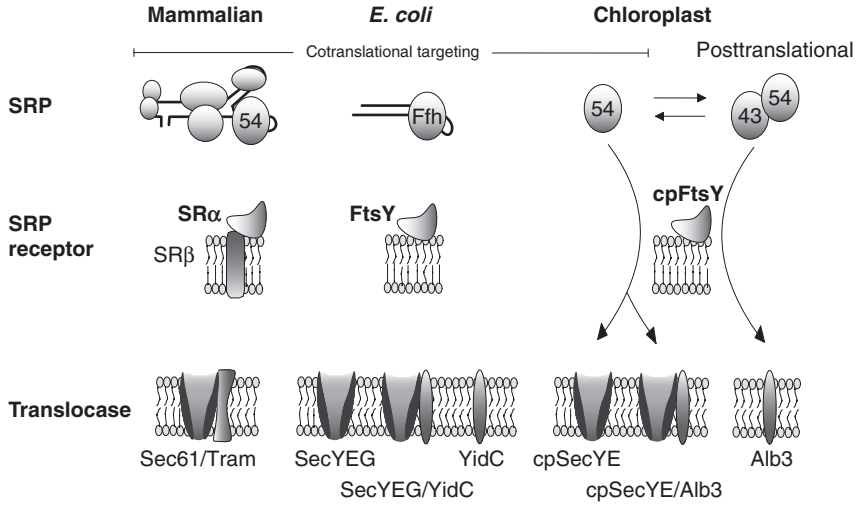



FIG. 19.2. Homologous components of the SRP pathway. Although there is considerable variation in signal recognition particles, many critical features of SRP, the SRP receptor, and translocase proteins are conserved. The cytosolic SRP in eukaryotes consists of an RNA moiety and six polypeptides, including the conserved GTPase SRP54. The bacterial SRP is composed of Ffh (54 homologue) and a smaller RNA moiety. cpSRP contains the conserved 54-kDa GTPase, but no RNA component. In addition, cpSRP contains a novel 43-kDa subunit that functions exclusively in posttranslational targeting. All SRP-targeting mechanisms rely on at least two GTPases, SRP54 noted above, and the receptor, SRα (FtsY in *E. coli* and cpFtsY in chloroplasts). In mammalian systems, the SRP receptor contains a second GTPase, the membrane protein SRβ. The translocase components of the SRP pathway are more diverse. In the ER, the translocation channel is formed by subunits of the Sec61 complex. The Sec61α and Sec61γ proteins of the Sec61 complex are similar in sequence to the SecY and SecE translocase components of the bacterial and chloroplast cotranslational targeting pathways. In bacteria and chloroplasts, the YidC/Alb3 families of proteins also function in membrane protein integration. For translocation into the lipid bilayer, association with TRAM is required. In the bacterial SRP pathway, conflicting reports of *in vitro* and *in vivo* protein requirements suggest a different translocase composition in relationship to the use of SRP. In chloroplasts, it appears that Alb3, like YidC, may have two distinct functions. cpSecYE may function in association with Alb3 during cotranslational integration, while a second pool of Alb3 functions independently of cpSecYE in posttranslational integration.

SRP RNA moiety in the chloroplast genomes of these organisms is consistent with the identification of amino acid substitutions at sites in cpSRP54 that would be expected to reduce their affinity for an SRP RNA [12, 13].

Further characterization of cpSRP led to the unexpected discovery of a novel 43-kDa subunit (cpSRP43) [14, 15]. Sequence analysis of the nuclear-encoded cpSRP43 predicts that nearly the entire protein is composed of known protein interaction domains, including three chromatin (CD)-binding

and four ankyrin (Ank)-binding domains (see [16, 17] for review, and see below). Unlike cpSRP54, the evolutionary origin of cpSRP43 remains uncertain in the absence of identifiable prokaryotic homologues. Biochemical assays demonstrated that cpSRP43 is required for transit complex formation, is a component of the cpSRP–LHCP transit complex, and is required for LHCP integration into isolated thylakoids [15]. It is notable that in stroma, cpSRP54 is found in two distinct pools, one associated with chloroplast ribosomes and another associated with cpSRP43 [15]. Only the cpSRP43-containing pool of cpSRP54, which appears to be a heterodimer with one copy of each subunit [12], exhibits the ability to bind full-length LHCPs and reconstitute thylakoid integration of LHCP. This is indicative of a specialized role for cpSRP43 in posttranslational cpSRP functions. This observation is also consistent with a dual function of cpSRP54 where the ribosome-associated pool of cpSRP54 is used for cotranslational routing of chloroplast-synthesized proteins (see below). High performance liquid chromatography analysis of recombinant cpSRP failed to detect GTP, suggesting that (as is the case for other SRPs) cpSRP is stable in the nucleotide-free form [12].

### B. A cpSRP Receptor Homologue Is Required for cpSRP-Based Protein Targeting to the Thylakoid Membrane

While it is possible that cpSRP serves simply as a chaperone to prevent the aggregation of LHC polypeptides as they traverse the stroma *en route* to the thylakoid, the need for GTP hydrolysis in LHCP integration supports the function of an SRP receptor at the thylakoid, and hence a membrane targeting role for cpSRP. Studies of protein targeting by cytosolic SRPs have shown that SRP receptors function at the target membrane to bind SRP54/Ffh, thereby delivering ribosome-bound targeting substrates to the membrane where the substrate can be released to transporters for cotranslational translocation into or across the lipid bilayer. The SRP receptor (SR) at the ER is composed of two GTPases, an integral β-subunit (SRβ) and a peripheral α-subunit (SRα) that interacts directly with SRP54 (Figure 19.2, for review see [10]). GTP hydrolysis by both SRα and SRP54 is used to release SRP from SR for subsequent rounds of targeting. SRP receptor function in *E. coli* is mediated by the SRα homologue FtsY, a GTPase that partitions between the soluble and membrane phases owing in part to its affinity for phospholipids [18]. Consistent with the role of GTP hydrolysis in releasing cpSRP from a thylakoid receptor, a chloroplast homologue of SRα and FtsY (cpFtsY) was identified in *Arabidopsis* by sequence homology. Antibodies to cpFtsY inhibited LHCP integration, establishing the participation of FtsY in posttranslational targeting by cpSRP [19]. Moreover, reconstitution of LHCP integration into isolated thylakoids requires cpFtsY along with cpSRP54, 43, and GTP

[20, 21]. Similar to *E. coli* FtsY, cpFtsY partitions between a soluble and a membrane-bound phase [20] and contains a conserved NG domain that is used in GTP binding and hydrolysis [19]. An *a*cidic (A) domain at the N-terminus of *E. coli* FtsY is absent in cpFtsY. However, nearly all of the A-domain is dispensable for FtsY function in *E. coli* [22] and numerous FtsY homologues in other prokaryotes lack an A-domain [23].

### C. LHCP Integration Appears Independent of Thylakoid Sec and TAT Transport Pathways

For the ER and bacterial systems, cotranslational insertion of SRP-targeted membrane proteins utilizes a homologous Sec translocase. In bacteria, the Sec translocase along with the SecA ATPase also supports transport of proteins across the lipid bilayer to the periplasm. Hence, both SRP- and SecA-dependent substrates can use SecYEG [24]. A Sec translocase in thylakoid membranes is composed minimally of SecY and SecE homologues (cpSecY and cpSecE, respectively) and functions along with the SecA homologue (cpSecA) to transport a subset of nuclear-encoded lumen-resident proteins across the lipid bilayer in an ATP-dependent mechanism (for review see [25]). Based on homology to bacterial protein sorting mechanisms, it would be expected that the cpSec translocase serves as the entry point for cpSRP-targeted proteins to enter the lipid bilayer and for cpSecA-dependent proteins to be transported across the bilayer into the lumen. While convergence at the cpSec translocase may hold true for cotranslational targeting by cpSRP, Mori and Cline [26] showed that antibody binding to cpSecY has no influence on LHCP integration despite inhibiting transport of lumen-resident proteins that require cpSecA. The same study also showed that LHCP integration was insensitive to antibodies that bind and inhibit the function of TAT translocase components. The results of these studies along with the fact that LHCP integration is sensitive to protease pretreatment of thylakoids [27] suggested that LHCP integration may utilize a protein-mediated insertion mechanism distinct from the TAT and Sec transporters.

### D. The Oxa1/Alb3/YidC Family Functions in LHCP Integration

Work in mitochondria, which lack Sec and SRP components, showed that a subset of mitochondrial- and nuclear-encoded inner membrane proteins rely on the function of Oxa1p for co- and posttranslational polypeptide integration from the matrix side [28]. Chloroplast and bacterial homologues

of Oxa1p, Alb3 and YidC were identified by sequence homology (see [29] for review). Consistent with the chlorophyll-less phenotype of nuclear insertion mutants that prevent Alb3 accumulation in *Arabidopsis* [30], Moore *et al.* [31] showed that antibody binding to Alb3 inhibits LHCP integration into isolated thylakoids without affecting protein transport across the thylakoid membrane by either the Sec or TAT translocation machinery.

Members of the Oxa1/Alb3/YidC family, which also includes Oxa2p in mitochondria and two Alb3 homologues in *Chlamydomonas* chloroplasts (see below), are polytopic membrane proteins that exhibit the greatest level of sequence similarity in their membrane-spanning regions. While Oxa1 and Alb3 each contains five hydrophobic membrane spans, YidC spans the cytoplasmic membrane six times owing to an additional nonconserved transmembrane region near its N-terminus [32]. As a result, both termini of YidC remain exposed on the cytosolic side of the membrane, whereas the N-terminus of both Oxa1p and Alb3 is translocated across the lipid bilayer and exposed to the mitochondrial inner membrane space or thylakoid lumen, respectively. YidC also differs from its organellar homologues in that it lacks an extended C-terminus. In mitochondria, the lengthy matrix-exposed C-terminus of Oxa1p is required for ribosome binding and cotranslational insertion of membrane proteins into the inner membrane, whereas Oxa2p lacks a ribosome-binding domain [33, 34]. The functional role of the stroma-exposed C-terminal extension in Alb3 is not known, but it is required for Alb3 to interact with cpSecY in a yeast split ubiquitin system [35].

Cross-linking studies have demonstrated that Alb3 is a nearest neighbor of cpSecY in the thylakoid [36]. This is consistent with studies in *E. coli* where a pool of YidC is found associated with the Sec translocase and appears to function in lateral movement of membrane-spanning domains into the lipid bilayer from the Sec complex during cotranslational integration [37, 38]. However, YidC plays a diverse role in membrane protein biogenesis. A Sec-independent pool of YidC is responsible for insertion of certain phage coat proteins as well as insertion of the $F_0$-subunit of the $F_1F_0$ ATP synthase by a mechanism that has been reconstituted with YidC proteoliposomes in the absence of SRP and Sec components [39]. Beyond its role in membrane protein insertion, YidC has also been shown to function as a membrane chaperone to promote proper folding and stability of LacY that is integrated in a YidC-independent manner [40]. Based on the fact that Alb3 can replace the function of YidC in *E. coli*, it is likely that Alb3 plays a similarly diverse role in the biogenesis of thylakoid membrane proteins, which is further supported by results of biochemical and genetic studies ([41] and see below).

### E. Posttranslational Binding to cpSRP Is Linked to an Alb3 Requirement for Integration

In *E. coli*, the Sec-independent mechanism for protein insertion still requires the function of YidC [42]. Hence, it would be expected that proteins thought to integrate spontaneously into thylakoids might actually rely on Alb3. Surprisingly, Robinson's group showed that only LHC polypeptides exhibit an Alb3 requirement for integration, which correlates with a strict requirement for cpSRP [43]. Other membrane proteins examined that are capable of posttranslational integration by an SRP-independent mechanism showed no requirement for functional Alb3. These data provided evidence that the cpSRP–LHCP transit complex may be targeted to the Alb3 translocase via interaction of cpSRP/cpFtsY with Alb3, but the relative dependence of specific LHCPs on cpSRP and cpFtsY *in vivo* differs (see below).

## V. Steps in the Posttranslational SRP Targeting Pathway

### A. cpSRP Assembly

#### *1. Posttranslational Targeting by cpSRP Requires Formation of a cpSRP54/cpSRP43 Dimer*

cpSRP54 is present in two pools in the chloroplast, a ribosome-associated, cotranslationally active form and a cpSRP43-associated, posttranslationally active form [6, 11, 15]. The interaction of cpSRP43 with cpSRP54 defines the targeting activity of cpSRP54 in the posttranslational pathway. A recombinant cpSRP43/cpSRP54 dimer supports LHCP integration, while the removal of either component prevents LHCP integration [12, 21]. Hence, assembly of the cpSRP43/cpSRP54 dimer is a prerequisite for cpSRP functions in posttranslational transport pathways.

cpSRP43 is an elongated molecule composed almost entirely of known protein interaction domains, an N-terminal CD (CD1) followed by four Ank repeat regions and two closely spaced CDs at the C-terminus (CD2 and CD3) ([12, 44, 45], Figure 19.3). Early work utilizing a yeast-two-hybrid system suggested that the two C-terminal CDs working together are responsible for cpSRP43 binding to cpSRP54 [44]. Subsequent analysis of glutathione *S*-transferase-fused CDs and domain deletions suggested that the central CD (CD2) alone is responsible for binding cpSPR54 [45]. Further analysis by isothermal titration calorimetry confirmed that CD2, and not the N- or C-terminal CDs, is involved in binding of cpSRP54 [46]. However, recent quantitative binding comparisons of CD2 containing constructs and full-length cpSRP43 suggest that other regions of cpSRP43 may also play a role in the binding of cpSRP43 to cpSRP54 [47].

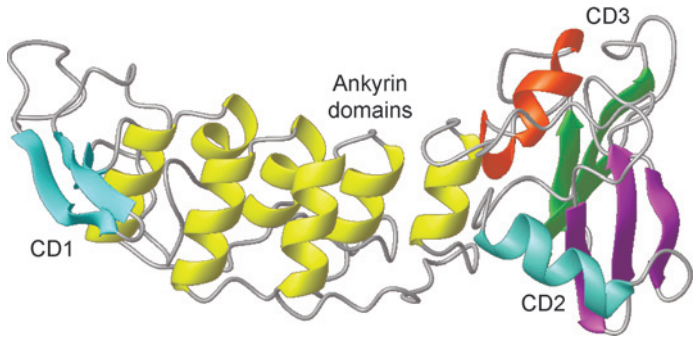



FIG. 19.3. cpSRP43 structural model. All three CDs of cpSRP43 contain a characteristic triple-stranded, antiparallel β-sheet as determined by NMR [46]. Several important structural differences are likely responsible for functional differences between the CDs (see below). CD1 lacks the characteristic C-terminal helix, while this helix is at notably different orientations in CD2 and CD3. Sequence analysis of the Ank region suggests the adoption of a typical canonical helix-loop-helix-β-hairpin-loop fold. Molecular modeling (courtesy of Kumar, T.K.S.), incorporating known NMR structures of the CDs and modeled structures of the Ank region have been used to generate a structural model for cpSRP43, which is consistent with its elongated biophysical characteristics [12].

CD structures are typically characterized by the presence of three antiparallel β-strands organized in a triple-stranded β-barrel structure along with a C-terminal α-helix. Structural comparisons of the three CDs in cpSRP43 revealed several important differences [46]. Although all three CDs contain a triple-stranded β-barrel structure, CD1 lacks the typical helical segment. Both CD2 and CD3 maintain a standard CD structure; however, the α-helix of CD2 is positioned perpendicular to the plane of the antiparallel β-sheet, while the α-helix of the C-terminal CD is in an unusual position parallel to the plane of the β-sheet (see Figure 19.3). A CD2 chimera in which the helix residues in CD2 were swapped for those in CD3 failed to bind cpSRP54 [46]. However, the CD2 α-helix did not impart cpSRP54 binding to the opposite chimera (Goforth and Henry, unpublished data). Together these data suggest that both residue identity and orientation may be important. Additional analysis of the residues within CD2 by mutagenesis and a pepscan approach have provided additional support for the importance of specific negatively charged amino acids in the α-helix as being important for the ability of cpSRP43 to coprecipitate critical regions of cpSRP54 [47].

An additional notable difference between the CDs of cpSRP43 is in the surface charge distribution. The surface charge distribution of the N-terminal CD is nearly zero, while in the C-terminal CD, the exterior of the β-sheet is neutral but the surface charge potential of the α-helix is positive [46]. In contrast, the surface charge potential of CD2 is highly negative owing to

the even surface distribution of acidic residues. This surface charge distribution seems especially relevant to cpSRP54 binding given the identification of essential, positively charged residues in cpSRP54 [46, 48].

cpSRP54 is composed of an N-terminal NG-domain responsible for a GTPase function and a C-terminal methionine-rich (M) domain [16]. Of particular note is the highly charged and extended nature of the C-terminal region of cpSRP54 as compared to the *E. coli* homologue Ffh [12]. A pepscan approach used to map the interactions between cpSRP43 and cpSRP54 using a peptide library of sequences generated for cpSRP54 from *Arabidopsis* revealed a large number of interactions with cpSPR43. These interacting peptides corresponded to areas in the G-domain near the consensus nucleotide-binding site, the flexible linker between the G- and M-domains, and portions of the M-domain including the C-terminal 26 residues [12]. Yeast-two-hybrid and coprecipitation data indicate that the M-domain of cpSRP54, and not the NG domains, is solely responsible for binding to cpSRP43 [44, 45]. Further research has narrowed identification of the interaction site for cpSRP43 to a unique 10-amino acid-long segment contained within the extended C-terminal region of the M-domain [48]. This peptide is characterized by the presence of a positively charged RRKRK sequence, and it seems likely that charge–charge interactions play an important role in interactions between this region of cpSRP54 and CD2. The presence of this sequence in cpSRP54, but not in Ffh, may explain the inability of Ffh to bind cpSRP43. It is interesting to question whether incorporation of this 10-amino acid sequence into Ffh would support binding of cpSRP43 to Ffh.

## B. Transit Complex Formation

### *1. cpSRP43 Binding to LHCP and cpSRP54 Coordinates Transit Complex Formation*

An investigation of the LHCP structural properties required for transit complex formation suggests that interactions between LHCP and cpSRP43 are required for cpSRP54 to bind hydrophobic sequences in LHCP. Formation of a cpSRP–LHCP transit complex requires two domains of LHCP, a hydrophobic domain and a unique SRP recognition element termed L18, which is composed of an 18-amino acid hydrophilic motif that resides between transmembrane domains 2 and 3 [49]. Consistent with the exclusive role of L18 in posttranslational cpSRP binding to LHCP family members, L18 functions as a cpSRP43-binding domain conserved among LHCP family members [50]. *Arabidopsis* LHCP family members are 50–83% identical compared to pea Lhcb (used to identify the L18 motif [49]) in

the region of the L18 motif [4]. Taken together, these data indicate that the L18-cpSRP43 interaction is used to recruit proteins to the posttranslational SRP sorting pathway in chloroplasts. This is largely supported in *Arabidopsis* mutants that lack cpSRP43 accumulation and exhibit a loss of LHC polypeptides, but accumulate normal levels of chloroplast-synthesized thylakoid proteins ([14] and see below). The possible role of the cpSRP43-L18 interaction in promoting cpSRP54 binding to hydrophobic sequences comes from studies in which L18 was fused to the N-terminus of bovine *p*re*p*ro*l*actin (PPL), a protein normally targeted to the ER by the cotranslational SRP targeting mechanism. While full-length PPL and L18-PPL showed almost no ability to bind cpSRP54 alone, the L18-PPL fusion bound to cpSRP43 formed transit complex with cpSRP54 [50]. Mutations to the hydrophobic PPL signal sequence that have been shown to inhibit signal sequence binding of PPL to cytosolic SRP54 in cotranslational assays also prevented formation of transit complex with L18-PPL without influencing interaction with cpSRP43. Taken together, these data support a model of transit complex formation in which posttranslational binding of hydrophobic substrates by the cpSRP54 M-domain only takes place following interaction of cpSRP43 with the L18 motif in the targeting substrate.

It is interesting that the structure of cytosolic SRP54 changes from a closed conformation to an extended conformation on binding to the ribosome, thereby exposing a hydrophobic groove in the M-domain to accommodate its interaction with hydrophobic signal sequences (for review see [51]). It is reasonable to expect that posttranslational binding of an LHCP hydrophobic region by cpSRP54 relies on a similar structural rearrangement of cpSRP54 within the M-domain. It is anticipated that the L18-cpSRP43 interaction may trigger the change in cpSRP54 so it can bind a nearby hydrophobic sequence in LHCP, similar to the function of the ribosome interaction with SRP54. It is noteworthy that cotranslational binding of signal peptide hydrophobic domains to all SRP54s requires a threshold of hydrophobicity [52–55]. When segments of LHCP were presented as ribosome nascent chains, efficient cross-linking of LHCP nascent chains to cpSRP54 or mammalian SRP54 was restricted to constructs that exposed the C-terminal membrane-spanning region, the most hydrophobic of the three membrane-spanning domains in LHCP [52]. Similar results were obtained by peptide scanning; only LHCP peptides containing the third membrane span showed interaction with cpSRP54 [12]. A minimal LHCP composed of the L18 motif and the third membrane span is sufficient to form transit complex with cpSRP [49]. Based on the fact that recombinant cpSRP is a heterodimer and the size of the cpSRP–LHCP transit complex is ~120 kDa, it is likely that the stoichiometry of cpSRP54, cpSRP43, and LHCP in transit complex is 1:1:1, but this remains to be firmly established.

Since transit complex appears to be the form of LHCP targeted to the thylakoid membrane, it seems likely by comparison to cotranslational targeting models [56] that cpSRP54 in transit complex is in a GTP-bound state (Figure 19.4). GTP is not required for transit complex formation, nor is it required for cotranslational binding of mammalian SRP54 to signal sequences. But a strict conservation of the GTP binding and hydrolysis cycle implies that affinity of cpSRP54 for GTP will be increased by one or more interactions that lead to transit complex formation in order to ready cpSRP54 for binding to its receptor, cpFtsY. Since L18 binding to cpSRP43 appears to signal the arrival of targeting substrate, it is interesting to speculate that binding of the L18 motif to cpSRP43 could act as the signal to increase affinity of cpSRP54 for GTP. Results of yeast-two-hybrid studies indicate that L18 binding takes place in the Ank repeat region of cpSRP43, specifically with Ank1 [44]. This is consistent with the results of transit complex formation assays using cpSRP43 domain deletions; a minimal cpSRP43 still capable of forming transit complex with cpSRP54 and LHCP is composed of the Ank repeat domain along with CD2. Both CD1 and CD3 are dispensable for transit complex formation [45].

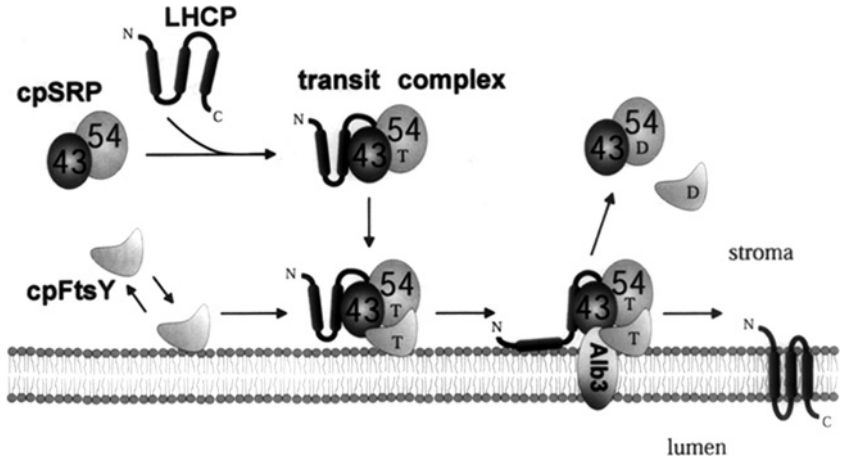



FIG. 19.4. Model of posttranslational cpSRP function in LHCP integration. cpSRP, composed of cpSRP54 (54) and cpSRP43 (43), interacts in solution with the substrate LHCP (black line, LHCP) to form a soluble intermediate termed transit complex. The cpSRP receptor, cpFtsY is located both in the soluble phase and at the thylakoid membrane. Transit complex arrives at the membrane loaded with substrate and interacts with cpFtsY. Interaction of cpSRP and cpFtsY likely promotes GTP binding by both cpSRP54 and cpFtsY (T, indicating GTP bound form). In the absence of accessible Alb3, the complex containing cpSRP, cpFtsY, and LHCP remains associated with the membrane until Alb3 is available. There the substrate is released to Alb3 for integration, and GTP hydrolysis (D, GDP-bound form) liberates cpSRP and cpFtsY for a successive round of targeting. See text for additional details.

## C. MEMBRANE EVENTS IN THE POSTTRANSLATIONAL CPSRP PATHWAY

### *1. Alb3 Is a Target for cpSRP–cpFtsY Binding at the Membrane*

The Alb3 requirement for stable integration of LHC polypeptides suggests that Alb3 is the membrane target for transit complex and that Alb3 functions to insert LHC polypeptides. Moreover, posttranslational targeting by cpSRP appears linked to the function of Alb3. SecY has also been shown to interact with at least a portion of the Alb3 in thylakoids [36]. Despite the fact that antibody binding to cpSecY does not affect LHCP integration [26], the inhibitory action of cpSecY antibody on transport of Sec pathway substrates could stem from inhibition of cpSecA binding to cpSecY rather than an inhibition of cpSecY function per se. Therefore, the membrane target for LHCPs may be cpSecY associated with Alb3. However, studies by Moore *et al.* [57] do not support a function of cpSecY in LHCP integration (see below). In either case, the mechanism for bringing LHC polypeptides to the target site at the membrane must include the action of cpFtsY, which likely binds cpSRP54 in a manner homologous to other SRP-targeting systems [58, 59]. Although cpFtsY partitions to both the stroma and thylakoid, there is no evidence that cpFtsY binds transit complex in solution. This is similar to *E. coli* where FtsY is found in both the soluble and membrane phase, but there is no requirement for FtsY to enter the soluble phase [60].

One approach to identify steps at the membrane involved in posttranslational targeting by cpSRP has been to use nonhydrolyzable GTP (GMP-PNP) to stabilize an interaction between cpSRP and cpFtsY. In the presence of thylakoids, a cpSRP–cpFtsY complex accumulates on the membrane in a GMP-PNP-dependent manner and copurifies as part of a larger complex containing Alb3 and cpSecY [57]. The cpSRP–cpFtsY complex occupies functional Alb3 translocation sites, demonstrated by decreased LHCP integration into thylakoids where this complex was stabilized prior to integration assays. Importantly, antibody binding to cpSecY removes cpSecY from membrane complexes containing cpSRP, cpFtsY, and Alb3 without inhibiting Alb3 activity. This argues that cpSecY is likely not part of the functional complex that is used to integrate LHCP. Interestingly, neither cpSRP43 nor LHCP is required to form a complex with Alb3, suggesting that cpSRP43 functions to link the substrate to the true targeting components, cpSRP54 and cpFtsY, which form the targeting–translocation interface with Alb3.

It is still uncertain if Alb3 function in LHC biogenesis is strictly limited to polypeptide insertion into the lipid bilayer. Since chlorophyll is required for stable insertion of LHC polypeptides, it has been hypothesized that chlorophyll attachment may take place at the site of LHC polypeptide insertion

(see [4] for review). One possibility is that enzymes which function late in chlorophyll biosynthesis are closely associated with Alb3. It is also possible that an added function of Alb3 is to hold inserted LHC polypeptides in position to promote chlorophyll attachment and proper folding, thereby acting as a membrane chaperone [32]. In *Chlamydomonas*, Alb3.1 functions as a membrane chaperone to promote assembly of inserted reaction center protein D1 into PS II, but is not required for D1 insertion into the lipid bilayer [61]. Characterization of cpSRP/cpFtsY/Alb3 complexes that represent functional LHCP integration sites in the thylakoid may lead to the identity of other components necessary for stabilizing LHC polypeptides during or after their insertion into the lipid bilayer.

### 2. *Targeting to the Membrane by cpSRP/cpFtsY May Precede Binding to Alb3*

It is noteworthy that in studies of cpSRP–cpFtsY binding to the thylakoid membrane, antibody binding to Alb3 prevented cpSRP–cpFtsY interaction with Alb3, but did not prevent attachment of the cpSRP–cpFtsY complex to the membrane [57]. Hence, LHCP targeting to Alb3 by cpSRP/cpFtsY may involve a step at the membrane that precedes interaction with Alb3 (refer to Figure 19.4). This step could require binding to an unknown membrane protein that has affinity for the targeting components or be mediated by affinity of the targeting components for lipids. Presumably, interaction(s) between Alb3 and the targeting components is required to initiate release of LHCP from both cpSRP54 and cpSRP43 and then initiate GTP hydrolysis by cpSRP54/cpFtsY to promote their separation from each other. A translocon-sensing mechanism appears to function in SRP targeting to the ER and may be a central feature of all SRP-targeting mechanisms, albeit with evolutionary adaptations unique to the target membrane and targeting components.

### 3. *cpSRP43 Functions at a Step Following Transit Complex Formation*

Although cpSRP54 and cpFtsY are able to interact with Alb3 in the absence of cpSRP43, we cannot rule out the possibility that cpSRP43 is involved in membrane events critical to control the timing of substrate release or control some aspects of the GTP cycle central to posttranslational integration of LHC polypeptides. Indeed, addition of cpSRP43 lacking CD1 has been shown to raise GTP hydrolysis rates by cpSRP54/cpFtsY without affecting the ability of cpSRP43 to bind LHCP and cpSRP54, or support transit complex formation [45]. The fact that cpSRP43 lacking CD1 loses the ability to support LHCP integration into isolated thylakoids without affecting any of its known soluble phase activities points to a critical role of cpSRP43 in directing events at the membrane. In the ER system, the reciprocal stimulation by SRP54 and

SRα of each protein's GTPase activities is promoted by the 60S ribosomal subunit [62], which has led to the intriguing possibility that cpSRP43 may functionally replace the ribosome as an activator of the cpSRP54/cpFtsY GTPase cycle [45]. A better understanding of cpSRP43's role at the membrane will require a detailed understanding of the interplay between the GTP cycle and protein interactions, especially those involving cpSRP43 at the membrane.

## VI. An Overlapping Post- and Cotranslational Function of cpSRP/cpFtsY/Alb3

### A. Biochemical Evidence for a Cotranslational cpSRP-Targeting Pathway

A limited number of cotranslational substrates (i.e., chloroplast-synthesized proteins) have been examined in biochemical assays that reconstitute thylakoid localization, making it difficult to fully understand how specific soluble and membrane components are linked. Localization conditions have only been established successfully for one chloroplast-synthesized thylakoid protein, cytochrome *f*, using isolated thylakoids and a homologous chloroplast translation system [63]. Cytochrome *f* is a single membrane-spanning protein that is localized using a cleavable signal peptide. Localization studies conducted with isolated thylakoids have shown that cpSRP does not bind the cytochrome *f* signal sequence as it emerges from the ribosome. Rather, the signal peptide interacts with cpSecA and integration is ATP dependent, which implicates cpSecYE function in the cotranslational integration of cytochrome *f*. The possibility exists that Alb3 may be required to move the single membrane-spanning domain from the cpSecYE pore into the lipid bilayer, similar to the function of Sec-associated YidC in *E. coli* (see Figure 19.2).

#### *1. Chloroplast-Synthesized D1 Binds Cotranslationally to cpSRP54 Prior to Integration by cpSecY*

The PS II reaction center protein, D1, is a polytopic membrane protein that has been shown to form cross-links with cpSRP54 when presented as a nascent chain synthesized *in vitro* by chloroplast ribosomes [64]. No interaction of D1 was observed with cpSRP43, which is consistent with the finding that only cpSRP54 is found associated with chloroplast ribosomes and functions in the absence of cpSRP43 to support cotranslational activities. Recruitment of D1 to the thylakoid membranes of intact chloroplasts results in its cotranslational integration, which proceeds through distinct

pauses in translation. D1 translation intermediates can be found associated both with the ribosome and cpSecY [65] implicating cpSecYE function in the integration mechanism. Presumably, cpFtsY acts in recruiting cpSRP54-bound nascent D1 to the thylakoid, but a requirement for cpFtsY has not been established for targeting activities that lead to D1 association with cpSecY (see below).

### 2. *Alb3 Is a Chaperone for Assembly of Integrated D1*

Based on the function of YidC in clearing transmembrane domains from the cpSecY translocase into the lipid bilayer, it would be predicted that Alb3 is necessary for D1 insertion into the membrane. However, a mutant of Alb3.1 in *Chlamydomonas* that exhibits a reduction in PS II assembly accumulates properly integrated D1 in thylakoids and is free of cpSecY [61, 66, 67]. The finding that Alb3.1 associates with fully integrated D1 suggests that Alb3.1 is not required to move D1 transmembrane domains from the Sec translocation pore into the lipid bilayer. Rather, Alb3.1 functions as a membrane chaperone for D1 to facilitate its assembly into PS II (see below). This is consistent with the diverse range of activities exhibited by the YidC/Alb3/Oxa1 family of proteins in bacteria and mitochondria.

## B. Analysis of Mutants Lacking Components of the cpSRP Pathway

### 1. *CpSRP Mutants*

*Arabidopsis* mutants lacking cpSRP43 (*chaos* mutant) [14], cpSRP54 (*ffc* mutant) [68], or both cpSRP subunits (*chaos/ffc* double mutant) [69] were analyzed. A summary of the main features of the mutants is given in Table 19.1. The *chaos* mutant was characterized by a specific defect in LHCP biogenesis, supporting the specialized role of cpSRP43 in LHCP targeting. The plant had a chlorotic phenotype with a 50% reduction of chlorophyll content. Western blot analysis revealed that the amount of most analyzed LHCPs is significantly reduced, whereas no reduction of other proteins could be detected.

Plants lacking functional cpSRP54 differed from the *chaos* mutant in two aspects [68, 70]. First, the phenotype of these plants was most visible at the young seedling stage when the first true leaves are yellow. In contrast to the *chaos* mutant, the phenotype becomes less severe when the plant matures, as seen by a recovery of the first true leaves and by the observation that the older *ffc* leaves are greener than the young *ffc* leaves. Second, the young leaves of the *ffc* mutant showed a reduced level of the chloroplast-encoded PS I and PS II reaction center proteins in addition to a reduction of the

## TABLE 19.1

*Arabidopsis*, Maize, *Chlamydomonas*, or *Synechocystis* Mutants Lacking Components of cpSRP or Members of the YidC/Oxa1p/Alb3 Family

| Mutant | Species | Mutated Protein | Appearance | Chlorophyll Content | Ultrastructure | Affected Proteins | References |
|---|---|---|---|---|---|---|---|
| *ffc* (x-ray) | *A. thaliana* | cpSRP54 | Yellow first true leaves, virescent | Reduction by 75% in first true leaves | Fewer thylakoids (first true leaves) | Reduction of most LHCPs; PS I, PS II reduced | [68] |
| *chaos* (transposon) | *A. thaliana* | cpSRP43 | Chlorotic, all leaves pale green | 50% reduction | Normal | Reduction of most LHCPs; PS I, PS II normal | [14] |
| *ffc/chaos* | *A. thaliana* | cpSRP54/ cpSRP43 | All leaves more yellow than chaos | 85% reduction | Strong reduction of thylakoids | Strong reduction of most LHCPs; PS I, PS II reduced | [69] |
| *csr1* (transposon) | Maize | cpFtsY | Pale yellow-green, seedling lethal | 88% reduction | Largely reduced and unstacked thylakoids | Strong reduction of most LHCPs; all photosystem complexes reduced; cpSecY, Alb3, cpTatC not affected; targeting and insertion of D1 not affected | [77] |
| *alb3* (transposon) | *A. thaliana* | Alb3 | Albino, seedling lethal | 95% reduction | Very few unstacked thylakoid membranes | | [30] |
| *alb4* (RNAi) | *A. thaliana* | Alb4 | Normal appearance | | Large chloroplasts; less appressed grana stacks | | [79] |
| *ac29* | *C. reinhardtii* | Alb3.1 | Yellow; retarded growth under low light | 70% reduction | | Strong reduction of LHC I and LHC II; PS I, ATP synthase and cytochrome $b_6f$ complex not affected; Reduced PS II; assembly of D1 into PS II affected | [66] |
| *alb3.2* (RNAi) | *C. reinhardtii* | Alb3.2 | Cell death after a prolonged period | 25–50% reduction | Enlarged vacuoles | Reduction of PS I and PS II (~50–75%); less severe reduction of LHC II, ATP synthase, and cytochrome $b_6f$ complex not affected; increased Vipp1, Hsp70, Cdj2 | [67] |
| *Δslr1471*Ω (disruption by antibiotic cassette) | *Synechocystis* | Alb3 | (No complete segregation achievable) | 60% reduction of pigments | Severe damage of thylakoid morphology | | [81] |

same subset of LHCPs as the *chaos* mutant. These data support the overlapping roles of cpSRP43 and cpSRP54 in LHCP targeting and the additional involvement of cpSRP54 in targeting chloroplast-encoded thylakoid membrane proteins.

Interestingly, the phenotype of a double mutant containing no functional cpSRP54 or cpSRP43 was more severe than that of the single mutants [69]. The double mutant had pale yellow leaves at all stages of growth and the levels of almost all LHCPs were significantly more reduced than in the single mutants. As expected, the double mutant exhibited a reduction of the chloroplast-encoded PS I and PS II reaction center proteins and the abundance of the inner antennae proteins CP43 and CP47 was also diminished. As in the single mutants, no reduction was observed for PsbS, a thylakoid membrane protein capable of spontaneous insertion [71]. The stronger phenotype of the double mutant suggests that the individual cpSRP subunits are partially active in LHCP targeting *in vivo*, although *in vitro* experiments demonstrated that the cpSRP subunits are inactive individually [12, 15, 16, 21]. Alternatively, it is also possible that the loss of just one subunit of the cpSRP can be compensated in the single mutants by the upregulation of factors that are active in combination with the remaining cpSRP subunit.

It should be noted that the cpSRP mutants were not seedling lethal, indicating that these mutant plants contained some photosynthetic activities. This is reflected by the observation that even the double mutant *ffc/chaos* still contained significant amounts of some members of the LHCP family (for review see [17]). Therefore, it may be proposed that the mutant plants can either partially compensate for the loss of cpSRP or are able to transport some LHCPs by an alternative cpSRP-independent pathway. An adaptation to the loss of the cotranslational SRP pathway by a slowdown of growth, a reduction of protein synthesis, and an induction of heat-shock proteins was described for *Saccharomyces cerevisiae* [72]. In addition, it was shown that *Streptococcus mutans* could also survive extensive disruptions of the SRP pathway [73]. Several studies analyzing the biogenesis of thylakoid membranes indicate that vesicles budding from the envelope membrane transport lipids to the growing thylakoids (reviewed in [74–76]). It is tempting to speculate that these vesicles might also be used to move proteins to the thylakoid membrane, although there is no indication for such a transport mechanism in chloroplasts.

### 2. *cpFtsY Mutant*

The analysis of a maize mutant containing a *Mu* transposon in the cpFtsY gene (*chloroplast SRP receptor1*; *csr1)* provides insights into the role of cpFtsY *in vivo* [77] (Table 19.1). The null cpFtsY mutant *csr1-1* was

seedling lethal and died shortly after germination. The young developing seedlings showed a chlorophyll-deficient phenotype with an otherwise almost normal morphology (reduction of chlorophyll to ~12%). Mutant bundle sheath and mesophyll chloroplasts possess an altered ultrastructure with largely reduced and unstacked thylakoid membranes. Detailed analysis of the steady state levels of individual members of the LHCP family revealed a drastic reduction of most LHCPs and further studies demonstrate that isolated mutant plastids import LHCP but are incapable of inserting LHCP into the thylakoid membrane. These data point to a crucial role of cpFtsY in LHCP biogenesis and support the previously published *in vitro* experiments demonstrating that an anti-cpFtsY antibody inhibited the LHCP insertion into thylakoid membranes [19] and that cpFtsY is required for the *in vitro* reconstitution of LHCP integration [20]. Interestingly, the *csr1* mutant does not only show a defect in LHCP biogenesis but is also characterized by a reduced abundance of all photosynthetic complexes, whereas the amount of subunits of the thylakoid protein transport machinery is not altered (e.g., cpSecY, Alb3, cpTatC). The observed pleiotropic defects of the *csr1* mutant point to an essential role of cpFtsY in the biogenesis of the photosynthetic apparatus, and it is tempting to speculate that cpFtsY might be involved in the cotranslational targeting of the chloroplast-encoded photosynthetic subunits. However, Asakura *et al.* showed that the targeting and insertion of D1, the reaction center protein of PS II, into the thylakoid membrane is not affected in the *csr1* mutant [77]. In summary, current data clearly demonstrate an important role of cpFtsY in the biogenesis of the LHCs, but the role of cpFtsY in the transport of other thylakoid membrane proteins remains unclear [77].

### *3. Alb3 Mutants*

The *albino3* (*alb3)* null mutant of *Arabidopsis*, generated by transposon mutagenesis, showed a drastic albinotic phenotype and was seedling lethal [30, 78]. Detailed analysis of the mutant plants revealed a strong chlorophyll deficiency (5% of wild-type chlorophyll content), abnormal chloroplasts with very few thylakoid membranes, and almost no grana stacking [30]. The phenotype of the *alb3* mutant resembles that of the cpFtsY mutant rather than the phenotypes of the *ffc*, *chaos*, and *ffc/chaos* mutants, which were not seedling lethal and contained some photosynthetic competency. This suggests that the function of Alb3 extends beyond insertion of a subset of LHCPs and is required for the insertion or assembly of other thylakoid membrane proteins. Alternatively, it is also possible that *Arabidopsis* mutants lacking cpSRP exhibit a milder phenotype since they might be able to compensate for this defect by upregulation of stromal chaperones.

A second chloroplast homologue of the Alb3/Oxa1/YidC family, named Alb4, was identified in *Arabidopsis* [79]. Alb4 represents the Alb3/Oxa1/YidC domain of the mistakenly identified 110-kDa inner envelope protein Artemis [80]. Like Alb3, Alb4 is expressed in the green tissues of *Arabidopsis* and is located in the thylakoid membrane. The strong phenotype of the *alb3* null mutant indicates that Alb4 cannot compensate the loss of Alb3 which points to a specialized function of Alb4. *Arabidopsis* mutants with a strong reduction (~90%) of the Alb4 level are vital and do not have an apparent visual phenotype under normal growth conditions. However, analysis of the ultrastructure of the mutant chloroplasts revealed that the mutant plastids are larger, more spherical in appearance and the grana stacks within the mutant lines are less appressed than in the wild-type chloroplasts. These data indicate that Alb4 is required for proper chloroplast biogenesis but nothing is yet known about the precise function of Alb4.

The green algae *Chlamydomonas reinhardtii* also contains two Alb3 homologues, Alb3.1 and Alb3.2 [66, 67]. A *Chlamydomonas* mutant lacking Alb3.1 (*ac29* mutant) is mainly characterized by a reduction of the light-harvesting systems. In addition, the *alb3.1* mutant has a reduced amount of PS II, whereas the abundance of PS I, the cytochrome $b_6f$ complex, and the ATP synthase is not altered [66]. Interestingly, another study demonstrated that the *alb3.1* mutant shows a defect in the assembly of D1 into PS II, whereas the integration of D1 into the thylakoid membrane appeared to be normal [61]. These data indicate that Alb3.1 has dual functions and is involved in the integration of the LHCPs and in the assembly of functional PS II. However, it should be pointed out that the mutant is still able to grow photoautotrophically and therefore exhibits a much milder phenotype than the *alb3* mutant of *Arabidopsis*.

Alb3.1 and Alb3.2 are both closely related to Alb3, whereby Alb3.2 shows a slightly higher homology to Alb3. It was demonstrated that depletion of Alb3.2 by RNA interference resulted, after a prolonged period, in cell death indicating an essential function of Alb3.2 in *Chlamydomonas* [67]. The abundance of PS I and PS II was reduced by 50–75% in the RNAi mutant, whereas the amount of light-harvesting chlorophyll a/b complex II was less affected. Remarkably, the effects of the Alb3.2 depletion were not restricted to the plastid since the mutant exhibited a large increase in vacuolar size [67]. These data indicate that Alb3.2 is not only involved in the assembly of photosynthetic complexes but has additional, not yet clarified essential functions.

Interestingly, *Synechocystis* sp. PCC6803 contains just a single *Alb3* gene (slr1471) [81]. Analysis of slr1471 mutant cells revealed that the Alb3 homologue is essential for cell viability and that depletion of this protein

results in severely damaged thylakoid morphology [81]. These data indicate that Alb3 plays a general role in protein insertion in *Synechocystis* (as in other bacteria) and is possibly required for the biogenesis of both the plasma membrane and the thylakoid membrane. A summary of the described mutants is given in Table 19.1.

## VII. Conclusions and Outlook

Biochemical and genetic studies of the posttranslational cpSRP–Alb3 transport pathway, coupled with knowledge of homologous targeting/integration systems, have provided a general picture of the steps required to target LHCPs to the membrane where interaction of cpSRP/cpFtsY with Alb3 culminates in LHCP integration. These studies indicate that components have evolved specialized functions unique to both the membrane target and to the requirements for targeting a full-length substrate. Yet, certain fundamental aspects of the SRP-targeting mechanism remain intact, including the use of GTP binding and hydrolysis by conserved GTPases to control the timing of events in both the soluble and membrane phases of the transport pathway. Presumably, this level of control is aimed at ensuring substrate release from SRP only in the presence of an available translocase. A detailed understanding of the relationship between GTP binding and the dynamics of targeting component interactions, with each other and with Alb3, will be required to elucidate details of a targeting mechanism that is anticipated to ''sense'' availability of Alb3.

It is interesting that genetic studies point to a more irreplaceable role of cpSRP43, cpFtsY, and Alb3 in posttranslational integration; the absence of cpSRP54 is more easily compensated as evidenced by a less severe phenotype. One possibility is that cpSRP43 binding to LHCPs serves a nonreplaceable function *in vivo*, whereas alternative mechanisms, for example stromal chaperones, compensate for the absence of cpSRP54. This raises the possibility that the severity of the cpFtsY phenotype is related to the commitment of targeting substrates to the cpSRP pathway on binding of cpSRP54, where cpFtsY may now be required to release cpSRP54 from the substrate. In this scenario, proteins targeted by both post- and cotranslational cpSRP-targeting mechanisms would be affected, which is consistent with the cpFtsY mutant phenotype. Given the apparent indispensability of Alb3 for integration of LHCPs (and possibly other thylakoid membrane proteins) *in vivo*, a mechanism must operate *in vivo* to deliver LHCPs to Alb3 that is independent of cpSRP54. At present, little is known about how Alb3 interacts with LHCPs (or chloroplast-synthesized membrane proteins) to promote their integration into the membrane. Development of a

cotranslational assay that faithfully reconstitutes integration into isolated thylakoids for a wide range of substrates would help elucidate the diverse roles of Alb3 in protein insertion or as a chaperone.

## Note Added In Proof

Tzvetkova-Chevolleau *et al.* [84] found that Arabidopsis containing a mutation in *cpftsy* fails to accumulate cpFtsY protein and exhibits a severe chlorotic phenotype similar to that observed for the *cpsrp54/chaos* double mutant lacking both cpSRP54 and cpSRP43. Intriguingly both thylakoid biogenesis and LHCP accumulation are largely normal when both cpFtsY and cpSRP54 are absent (*cpftsy/cpsrp54* double mutant) yet cpSRP43 is present. The authors also demonstrate that cpSRP43, in the absence of cpSRP54 and cpFtsY, exhibits the ability to bind the Alb3 translocase. Taken together, these findings indicate that cpSRP43 has developed features to funtion independently of cpSRP54/cpFtsY in targeting LHCPs to the thylakoid memberane and serves as a component of an alternative pathway for targeting to the Alb3 translocase.

## Acknowledgments

The authors thank T.K.S. Kumar for sharing the modeled structure of cpSRP43. Work in the author's laboratories was supported by Department of Energy Grant BR15569 (R.H.), NIH Grant P20 RR15569 from the COBRE Program of the National Center for Research Resources (R.G.), and Deutsche Forschungsgemeinschaft Grant Number SFB 480 B11 (D.S.).

## References

1. Schatz, G., and Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science* 271:1519–1526.
2. Vothknecht, U.C., and Soll, J. (2005). Chloroplast membrane transport: interplay of prokaryotic and eukaryotic traits. *Gene* 354:99–109.
3. Kessler, F., and Schnell, D.J. (2006). The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. *Traffic* 7:248–257.
4. Cline, K. (2003). Biogenesis of green plant thylakoid membranes. *Adv Photosynth Respir* 13:353–372.
5. Payan, L.A., and Cline, K. (1991). A stromal protein factor maintains the solubility and insertion competence of an imported thylakoid membrane protein. *J Cell Biol* 112:603–613.
6. Franklin, A.E., and Hoffman, N.E. (1993). Characterization of a chloroplast homologue of the 54-kDa subunit of the signal recognition particle. *J Biol Chem* 268:22175–22180.

7. Hoffman, N.E., and Franklin, A.E. (1994). Evidence for a stromal GTP requirement for the integration of a chlorophyll a/b-binding polypeptide into thylakoid membranes. *Plant Physiol* 105:295–304.
8. Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001). The signal recognition particle. *Annu Rev Biochem* 70:755–775.
9. Luirink, J., and Sinning, I. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta* 1694:17–35.
10. Pool, M. (2005). Signal recognition particles in chloroplasts, bacteria, yeast and mammals. *Mol Membr Biol* 22:3–15.
11. Li, X., Henry, R., Yuan, J., Cline, K., and Hoffman, N.E. (1995). A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. *Proc Natl Acad Sci USA* 92:3789–3793.
12. Groves, M.R., Mant, A., Kuhn, A., Koch, J., Dubel, S., Robinson, C., and Sinning, I. (2001). Functional characterization of recombinant chloroplast signal recognition particle. *J Biol Chem* 276:27778–27786.
13. Rosenblad, M.A., and Samuelsson, T. (2004). Identification of chloroplast signal recognition particle RNA genes. *Plant Cell Physiol* 45:1633–1639.
14. Klimyuk, V.I., Persello-Cartieaux, F., Havaux, M., Contard-David, P., Schuenemann, D., Meiherhoff, K., Gouet, P., Jones, J.D., Hoffman, N.E., and Nussaume, L. (1999). A chromodomain protein encoded by the Arabidopsis CAO gene is a plant-specific component of the chloroplast signal recognition particle pathway that is involved in LHCP targeting. *Plant Cell* 11:87–100.
15. Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V.I., Jones, J.D.G., Nussaume, L., and Hoffman, N.E. (1998). A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc Natl Acad Sci USA* 95:10312–10316.
16. Eichacker, L.A., and Henry, R. (2001). Function of a chloroplast SRP in thylakoid protein export. *Biochim Biophys Acta* 1541:120–134.
17. Schuenemann, D. (2004). Structure and function of the chloroplast signal recognition particle. *Curr Genet* 44:295–304.
18. de Leeuw, E., te Kaat, K., Moser, C., Menestrina, G., Demel, R., de Kruijff, B., Oudega, B., Luirink, J., and Sinning, I. (2000). Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity. *EMBO J* 19:531–541.
19. Kogata, N., Nishio, K., Hirohashi, T., Kikuchi, S., and Nakai, M. (1999). Involvement of a chloroplast homologue of the signal recognition particle receptor protein, FtsY, in protein targeting to thylakoids. *FEBS Lett* 447:329–333.
20. Tu, C.J., Schuenemann, D., and Hoffman, N.E. (1999). Chloroplast FtsY, chloroplast signal recognition particle, and GTP are required to reconstitute the soluble phase of light-harvesting chlorophyll protein transport into thylakoid membranes. *J Biol Chem* 274:27219–27224.
21. Yuan, J., Kight, A., Goforth, R.L., Moore, M., Peterson, E.C., Sakon, J., and Henry, R. (2002). ATP stimulates signal recognition particle (SRP)/FtsY-supported protein integration in chloroplasts. *J Biol Chem* 277:32400–32404.
22. Eitan, A., and Bibi, E. (2004). The core *Escherichia coli* signal recognition particle receptor contains only the N and G domains of FtsY. *J Bacteriol* 186:2492–2494.
23. Gorodkin, J., Knudsen, B., Zwieb, C., and Samuelsson, T. (2001). SRPDB (signal recognition particle database). *Nucleic Acids Res* 29:169–170.
24. de Gier, J.W., and Luirink, J. (2001). Biogenesis of inner membrane proteins in *Escherichia coli*. *Mol Microbiol* 40:314–322.
25. Mori, H., and Cline, K. (2001). Post-translational protein translocation into thylakoids by the Sec and DeltapH-dependent pathways. *Biochim Biophys Acta* 1541:80–90.

26. Mori, H., Summer, E.J., Ma, X., and Cline, K. (1999). Component specificity for the thylakoidal sec and deltapH-dependent protein transport pathways. *J Cell Biol* 146:45–56.
27. Robinson, D., Karnauchov, I., Herrmann, R.G., Klosgen, R.B., and Robinson, C. (1996). Protease-sensitive thylakoidal import machinery for the Sec-, Delta-pH- and signal recognition particle-dependent protein targeting pathways, but not for CFoII integration. *Plant J* 10:149–155.
28. Stuart, R.A. (2002). Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. *Biochim Biophys Acta* 1592:79–87.
29. Yi, L., and Dalbey, R. (2005). Oxa1/Alb3/YidC system for insertion of membrane proteins in mitochondria, chloroplasts and bacteria (review). *Mol Membr Biol* 22:101–111.
30. Sundberg, E., Slagter, J.G., Fridborg, I., Cleary, S.P., Robinson, C., and Coupland, G. (1997). ALBINO3, an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. *Plant Cell* 9:717–730.
31. Moore, M., Harrison, M.S., Peterson, E.C., and Henry, R. (2000). Chloroplast oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. *J Biol Chem* 275:1529–1532.
32. Kuhn, A., Stuart, R., Henry, R., and Dalbey, R.E. (2003). The Alb3/Oxa1/YidC protein family: membrane-localized chaperones facilitating membrane protein insertion? *Trends Cell Biol* 13:510–516.
33. Szyrach, G., Ott, M., Bonnefoy, N., Neupert, W., and Herrmann, J.M. (2003). Ribosome binding to the Oxa1 complex facilitates co-translational protein insertion in mitochondria. *EMBO J* 22:6448–6457.
34. Funes, S., Nargang, F.E., Neupert, W., and Herrmann, J.M. (2004). The Oxa2 protein of Neurospora crassa plays a critical role in the biogenesis of cytochrome oxidase and defines a ubiquitous subbranch of the Oxa1/YidC/Alb3 protein family. *Mol Biol Cell* 15:1853–1861.
35. Pasch, J.C., Nickelsen, J., and Schuenemann, D. (2005). The yeast split-ubiquitin system to study chloroplast membrane protein interactions. *Appl Microbiol Biotechnol* 69:440–447.
36. Klostermann, E., Droste Gen Helling, I., Carde, J.P., and Schunemann, D. (2002). The thylakoid membrane protein ALB3 associates with the cpSecY-translocase in *Arabidopsis thaliana*. *Biochem J* 368:777–781.
37. Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000). YidC mediates membrane protein insertion in bacteria. *Nature* 406:637–641.
38. Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W., von Heijne, G., van der Does, C., Driessen, A.J., Oudega, B., and Luirink, J. (2000). YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J* 19:542–549.
39. van der Laan, M., Nouwen, N.P., and Driessen, A.J.M. (2005). YidC—an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr Opin Microbiol* 8:182–187.
40. Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004). Role of YidC in folding of polytopic membrane proteins. *J Cell Biol* 165:53–62.
41. Jiang, F., Yi, L., Moore, M., Chen, M., Rohl, T., Van Wijk, K.J., De Gier, J.W., Henry, R., and Dalbey, R.E. (2002). Chloroplast YidC homolog Albino3 can functionally complement the bacterial YidC depletion strain and promote membrane insertion of both bacterial and chloroplast thylakoid proteins. *J Biol Chem* 277:19281–19288.

42. Van Bloois, E., Haan, G.J., De Gier, J.W., Oudega, B., and Luirink, J. (2006). Distinct requirements for translocation of the N-tail and C-tail of the *Escherichia coli* inner membrane protein CyoA. *J Biol Chem* 281:10002–10009.
43. Mant, A., Woolhead, C.A., Moore, M., Henry, R., and Robinson, C. (2001). Insertion of PsaK into the thylakoid membrane in a ''Horseshoe'' conformation occurs in the absence of signal recognition particle, nucleoside triphosphates, or functional albino. *J Biol Chem* 276:36200–36206.
44. Jonas-Straube, E., Hutin, C., Hoffman, N.E., and Schunemann, D. (2001). Functional analysis of the protein-interacting domains of chloroplast SRP43. *J Biol Chem* 276:24654–24660.
45. Goforth, R.L., Peterson, E.C., Yuan, J., Moore, M.J., Kight, A.D., Lohse, M.B., Sakon, J., and Henry, R.L. (2004). Regulation of the GTPase cycle in post-translational signal recognition particle-based protein targeting involves cpSRP43. *J Biol Chem* 279:43077–43084.
46. Sivaraja, V., Kumar, T.K.S., Leena, P.S.T., Chang, A.N., Vidya, C., Goforth, R.L., Rajalingam, D., Arvind, K., Ye, J.L., Chou, J., Henry, R., and Yu, C. (2005). Three-dimensional solution structures of the chromodomains of cpSRP43. *J Biol Chem* 280:41465–41471.
47. Hermkes, R., Funke, S., Richter, C., Kuhlmann, J., and Schuenemann, D. (2006). The a-helix of the second chromodomain of the 43kDa subunit of the chloroplast signal recognition particle facilitates binding to the 54kDa subunit. *FEBS Lett* 580:3107–3111.
48. Funke, S., Knechten, T., Ollesch, J., and Schuenemann, D. (2005). A unique sequence motif in the 54-kDa subunit of the chloroplast signal recognition particle mediates binding to the 43-kDa subunit. *J Biol Chem* 280:8912–8917.
49. DeLille, J., Peterson, E.C., Johnson, T., Moore, M., Kight, A., and Henry, R. (2000). A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. *Proc Natl Acad Sci USA* 97:1926–1931.
50. Tu, C.J., Peterson, E.C., Henry, R., and Hoffman, N.E. (2000). The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J Biol Chem* 275:13187–13190.
51. Halic, M., and Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. *Curr Opin Struct Biol* 15:116–125.
52. High, S., Henry, R., Mould, R.M., Valent, Q., Meacock, S., Cline, K., Gray, J.C., and Luirink, J. (1997). Chloroplast SRP54 interacts with a specific subset of thylakoid precursor proteins. *J Biol Chem* 272:11622–11628.
53. Lee, H.C., and Bernstein, H.D. (2001). The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc Natl Acad Sci USA* 98:3471–3476.
54. Valent, Q.A., Kendall, D.A., High, S., Kusters, R., Oudega, B., and Luirink, J. (1995). Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides. *EMBO J* 14:5494–5505.
55. Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* 134:269–278.
56. Wild, K., Rosendal, K.R., and Sinning, I. (2004). A structural step into the SRP cycle. *Mol Microbiol* 53:357–363.
57. Moore, M., Goforth, R.L., Mori, H., and Henry, R. (2003). Functional interaction of chloroplast SRP/FtsY with the ALB3 translocase in thylakoids: substrate not required. *J Cell Biol* 162:1245–1254.
58. Focia, P.J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004). Heterodimeric GTPase core of the SRP targeting complex. *Science* 303:373–377.

59. Egea, P.F., Shan, S.O., Napetschnig, J., Savage, D.F., Walter, P., and Stroud, R.M. (2004). Substrate twinning activates the signal recognition particle and its receptor. *Nature* 427:215–221.
60. Zelazny, A., Seluanov, A., Cooper, A., and Bibi, E. (1997). The NG domain of the prokaryotic signal recognition particle receptor, FtsY, is fully functional when fused to an unrelated integral membrane polypeptide. *Proc Natl Acad Sci USA* 94:6025–6029.
61. Ossenbuehl, F., Goehre, V., Meurer, J., Krieger-Liszkay, A., Rochaix, J.D., and Eichacker, L.A. (2004). Efficient assembly of photosystem II in Chlamydomonas reinhardtii requires Alb3.p, a homolog of arabidopsis ALBINO3. *Plant Cell* 16:1790–1800.
62. Mandon, E.C., Jiang, Y., and Gilmore, R. (2003). Dual recognition of the ribosome and the signal recognition particle by the SRP receptor during protein targeting to the endoplasmic reticulum. *J Cell Biol* 162:575–585.
63. Rohl, T., and van Wijk, K.J. (2001). *In vitro* reconstitution of insertion and processing of cytochrome f in a homologous chloroplast translation system. *J Biol Chem* 276:35465–35472.
64. Nilsson, R., and van Wijk, K.J. (2002). Transient interaction of cpSRP54 with elongating nascent chains of the chloroplast-encoded D1 protein; 'cpSRP54 caught in the act.' *FEBS Lett* 524:127–133.
65. Zhang, L., Paakkarinen, V., Suorsa, M., and Aro, E.M. (2001). A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. *J Biol Chem* 276:37809–37814.
66. Bellafiore, S., Ferris, P., Naver, H., Gohre, V., and Rochaix, J.D. (2002). Loss of Albino3 leads to the specific depletion of the light-harvesting system. *Plant Cell* 14:2303–2314.
67. Gohre, V., Ossenbuhl, F., Crevecoeur, M., Eichacker, L.A., and Rochaix, J.D. (2006). One of two Alb3 proteins is essential for the assembly of the photosystems and for cell survival in Chlamydomonas. *Plant Cell* 18:1454–1466.
68. Amin, P., Sy, D.A., Pilgrim, M.L., Parry, D.H., Nussaume, L., and Hoffman, N.E. (1999). Arabidopsis mutants lacking the 43- and 54-kilodalton subunits of the chloroplast signal recognition particle have distinct phenotypes. *Plant Physiol* 121:61–70.
69. Hutin, C., Havaux, M., Carde, J.P., Kloppstech, K., Meiherhoff, K., Hoffman, N.E., and Nussaume, L. (2002). Double mutation cpSRP43/cpSRP54 is necessary to abolish the cpSRP pathway required for thylakoid targeting of the light-harvesting chlorophyll proteins. *Plant J* 29:531–543.
70. Pilgrim, M.L., van Wijk, K.J., Parry, D.H., Sy, D.A., and Hoffman, N.E. (1998). Expression of a dominant negative form of cpSRP54 inhibits chloroplast biogenesis in Arabidopsis. *Plant J* 13:177–186.
71. Kim, S.J., Jansson, S., Hoffman, N.E., Robinson, C., and Mant, A. (1999). Distinct ''assisted'' and ''spontaneous'' mechanisms for the insertion of polytopic chlorophyll-binding proteins into the thylakoid membrane. *J Biol Chem* 274:4715–4721.
72. Mutka, S.C., and Walter, P. (2001). Multifaceted physiological response allows yeast to adapt to the loss of the signal recognition particle-dependent protein-targeting pathway. *Mol Biol Cell* 12:577–588.
73. Hasona, A., Crowley, P.J., Levesque, C.M., Mair, R.W., Cvitkovitch, D.G., Bleiweis, A.S., and Brady, L.J. (2005). Streptococcal viability and diminished stress tolerance in mutants lacking the signal recognition particle pathway or YidC2. *Proc Natl Acad Sci USA* 102:17466–17471.
74. Hoober, J.K., and Eggink, L.L. (1999). Assembly of light-harvesting complex II and biogenesis of thylakoid membranes in chloroplasts. *Photosynth Res* 61:197–215.
75. Vothknecht, U.C., and Westhoff, P. (2001). Biogenesis and origin of thylakoid membranes. *Biochim Biophys Acta* 1541:91–101.
76. Benning, C., Xu, C., and Awai, K. (2006). Non-vesicular and vesicular lipid trafficking involving plastids. *Curr Opin Plant Biol* 9:241–247.

77. Asakura, Y., Hirohashi, T., Kikuchi, S., Belcher, S., Osborne, E., Yano, S., Terashima, I., Barkan, A., and Nakai, M. (2004). Maize mutants lacking chloroplast FtsY exhibit pleiotropic defects in the biogenesis of thylakoid membranes. *Plant Cell* 16:201–214.
78. Long, D., Martin, M., Sundberg, E., Swinburne, J., Puangsomlee, P., and Coupland, G. (1993). The maize transposable element system Ac/Ds as a mutagen in Arabidopsis: identification of an albino mutation induced by Ds insertion. *Proc Natl Acad Sci USA* 90:10370–10374.
79. Gerdes, L., Bals, T., Klostermann, E., Karl, M., Philippar, K., Huenken, M., Soll, J., and Schuenemann, D. (2006). A second thylakoid membrane-localized Alb3/OxaI/YidC homologue is involved in proper chloroplast biogenesis in *Arabidopsis thaliana*. *J Biol Chem* 281:16632–16642.
80. Fulgosi, H., Gerdes, L., Westphal, S., Glockmann, C., and Soll, J. (2002). Cell and chloroplast division requires ARTEMIS. *Proc Natl Acad Sci USA* 99:11501–11506.
81. Spence, E., Bailey, S., Nenninger, A., Moller, S.G., and Robinson, C. (2004). A homolog of Albino3/OxaI is essential for thylakoid biogenesis in the cyanobacterium *Synechocystis* sp. PCC6803. *J Biol Chem* 279:55792–55800.
82. Theg, S.M., Cline, K., Finazzi, G., and Wollman, F.A. (2005). The energetics of the chloroplast Tat protein transport pathway revisited. *Trends Plant Sci* 10:153–154.
83. Di Cola, A., Klostermann, E., and Robinson, C. (2005). The complexity of pathways for protein import into thylakoids: it's not easy being green. *Biochem Soc Trans* 33:1024–1027.
84. Tzvetkova-Chevolleau, T., Hutin, C., Noël, L.D., Goforth, R.L., Carde, J-P., Caffarri, S., Sinning, I., Groves, M., Teulon, J.-M., Hoffman, N.E., Henry, R., Havaux, M., and Nussaume, L. (2007). Canonical SRP components can be bypassed for post-translational protein targeting in chloroplasts. *Plant Cell* doi/10.1105/tpc.106.048959

This page intentionally left blank

# Part V

*Crossing Peroxisomal Membranes*

This page intentionally left blank

# 20

# *The Role of Shuttling Targeting Signal Receptors and Heat-Shock Proteins in Peroxisomal Matrix Protein Import*

SEBASTIEN LEON[a] • SURESH SUBRAMANI[b]

[a]*Institut Jacques Monod-CNRS*
*2 Place Jussieu*
*75251 Paris Cedex 05*
*France*

[b]*Section of Molecular Biology*
*Division of Biological Sciences*
*University of California, San Diego*
*La Jolla, CA 92093-0322, USA*

## I. Catalytic Machines Involved in Peroxisomal Matrix Protein Import

Peroxisome biogenesis requires the action of about 32 *PEX* genes encoding a family of proteins known as peroxins [1]. These are distributed in the cytosol, peroxisomal membrane, or peroxisome lumen. While a number of these proteins, particularly components of a peroxisome membrane-associated complex known as the importomer [2], regulate the entry of cargoes and cargo receptors into and out of peroxisomes, a number of proteins act catalytically to facilitate the import of proteins into the peroxisome matrix. This chapter will focus on two sets of proteins that act in this fashion.



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25020-6

1. The peroxisome-targeting signal (PTS) receptors that repeatedly bind cargo in the cytosol ferry the cargo into the peroxisome lumen, and then recycle back to the cytosol for another round of import [3]. When recycling of these receptors is compromised, enzymes involved in a peroxisomal pathway known as receptor accumulation and degradation in the absence of recycling (RADAR) function to keep the peroxisome membrane clear of cargo-free receptors [3]. These will also be described briefly.
2. The Hsp70 class of proteins that undergoes cycles of ATP binding and hydrolysis.

Other enzymes required for peroxisomal matrix protein import include those, such as Pex4p, involved in protein ubiquitylation, and the AAA ATPases, Pex1 and Pex6, that associate with each other and are needed for receptor recycling from peroxisomes to the cytosol—these are reviewed separately in this volume.

## II. Components Involved in Peroxisomal Matrix Protein Import

To set the stage for the discussion of these catalytic machines, we first need to understand the peroxisome biogenesis machinery itself and the general mechanism by which proteins are sorted to the peroxisome matrix.

Most proteins destined for the peroxisome matrix possess one or more PTSs. A few peroxisomal matrix proteins lack a PTS but enter the matrix simply by association with other protein partners that do have a PTS, which is a consequence of the fact that folded, oligomeric and cofactor-bound proteins can be transported into the peroxisome lumen [3]. Two classes of conserved PTSs have been described—a C-terminal tripeptide or PTS1 and an N-terminal or internal nona-peptide, named PTS2. Following the synthesis of peroxisomal matrix proteins in the cytosol, these PTSs are recognized by specific receptors/coreceptors. The Pex5 protein is the PTS1 receptor [4–7], and the Pex7 protein is the PTS2 receptor in yeasts, plants, and mammals [8–15]. However, the PTS2 pathway of import requires an auxiliary protein, or coreceptor, to transport cargoes into the matrix. In yeasts, members of the Pex20 family of proteins (Pex18 or Pex21 in *Saccharomyces cerevisiae* and Pex20 in *Neurospora crassa*, *Pichia pastoris*, and *Yarrowia lipolytica*) are also necessary for the PTS2 import pathway [16–20], whereas in plants and mammals, a long isoform of Pex5, called Pex5L, serves as the auxiliary protein for the PTS2 pathway [15, 21].

The complexes formed between cargo and the PTS coreceptors and/or receptors then interact at the peroxisomal membrane with a docking

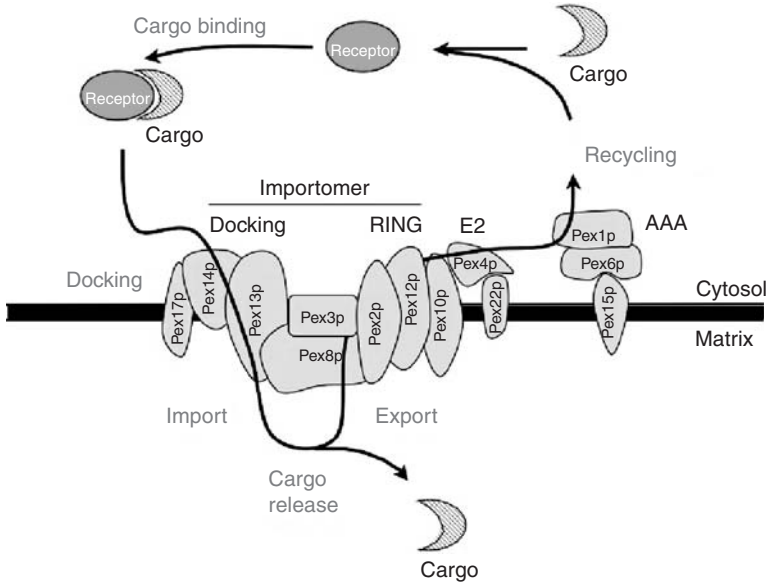



FIG. 20.1. Steps in the extended shuttle cycle of the PTS receptors during peroxisomal matrix protein import. See text for details.

subcomplex, comprised in yeasts of the peroxisomal membrane or membrane-associated peroxins, Pex3, Pex8, Pex13, and Pex14 (Figure 20.1). This docking subcomplex likely serves as the translocon for the receptor/cargo complexes [3] and associates with a second subcomplex in the peroxisome membrane, the really interesting new gene (RING) subcomplex, composed of three RING-domain proteins, Pex2, Pex10, and Pex12 [22, 23]. The docking and RING subcomplexes have been reported in different yeasts to be bridged either by Pex3 or by Pex8 to form a larger complex known as the importomer, all of whose individual constituents are necessary for import of most PTS1- and PTS2-containing proteins [22, 23].

The importomer is also in proximity, if not association, with a receptor recycling machinery [17, 24, 25], comprised in yeasts of the E2-like ubiquitin-conjugating enzyme, Pex4, anchored to the peroxisome membrane via Pex22, and two AAA ATPases, Pex1 and Pex6, that associate with the peroxisome membrane (via Pex15 in *S. cerevisiae* and PEX26 in mammals) (Figure 20.1).

There is evidence from epistasis and biochemical analyses that the docking, RING, and recycling subcomplexes act sequentially during the matrix protein import cycle [17, 26].

### A. Receptor Shuttling During Peroxisomal Matrix Protein Import

The PTS receptors function in multiple rounds of cargo import into peroxisomes [24, 25, 27]. They do so by binding cargo in the cytosol, ferrying it to the peroxisome lumen, releasing cargo in the peroxisome matrix, and recycling back to the cytosol. This cycle of receptor dynamics is referred to as the extended shuttle [27, 28]. This mechanism allows each molecule of receptor to participate in several rounds of cargo import into peroxisomes. We consider first the experimental evidence for the extended shuttle model of receptor dynamics and then outline features of the specific steps in receptor shuttling.

Early work on the subcellular location of the PTS1 receptor, Pex5, was confusing because in different organisms it was found to be either cytosolic, or mostly cytosolic and partially peroxisomal, or inside peroxisomes. These could have been attributed to organism-specific variations. However, when the *S. cerevisiae* PTS2 receptor was identified by two laboratories, it was very puzzling that one group reported Pex7 to be cytosolic, while the other found it to be intraperoxisomal [10, 13, 14] (only later it was found that the predominantly intraperoxisomal localizaion of Pex7 was caused by the attachment of a C-terminal tag used to localize the protein [29]). These discrepancies suggested the possibility of a shuttling receptor [10, 30]. It was hypothesized that the use of tagged proteins to follow their subcellular location might alter the steady state distributions of these proteins among the compartments between which they shuttle.

More careful experiments with human PEX5 revealed a dual location with most of the protein being cytosolic but some was peroxisome-associated [4, 7]. The ability to shift the equilibrium between these two populations, either by modulation of conditions used for peroxisomal protein import or the use of mutants, suggested that there might be a dynamic movement of PEX5 from the cytosol to the peroxisomes. For example, under low temperature or ATP-depletion conditions that affect matrix protein import, more PEX5 was peroxisome-asssociated, and the peroxisome-associated pool was reduced on raising the temperature or on readdition of ATP [31]. Similar perturbations of the mammalian PEX5 population in the cytosol versus the peroxisomes were observed in certain *pex* mutants affecting peroxisome biogenesis [31]. Intraperoxisomal accumulation of Pex5 was also observed in the *Hansenula polymorpha pex4* mutant [32]. As a result of these data, the idea of a cycling receptor was given more credence, but these early studies assumed implicitly that the receptor ferried cargo to the peroxisomal translocon, released cargo there, and returned directly to the cytosol for another round of import (simple shuttle) [10]. Although the entry of the receptor into peroxisomes had been suggested [30], experimental data supporting this notion was lacking.

Evidence that PEX5 uses instead an extended shuttle, where it goes in and out of peroxisomes during the import cycle, came from the work of Dammai and Subramani [27]. In this study, a modified version of PEX5 was processed by a peroxisome-specific protease, before returning to the cytosol, indicating that PEX5 can sample, if not enter, the peroxisome matrix. The study also suggested that specific *cis*-acting sequences and *trans*-acting proteins may be necessary for PEX5 export from peroxisomes and recycling to the cytosol—predictions now supported by experimental data (see below).

Experiments using *in vitro* systems have addressed the requirements for the association of PEX5 with peroxisomes, as well as for its recycling to the cytosol [24, 33, 34]. These experiments also shed light on the steps in the PEX5 extended shuttle cycle. The existence of several populations of PEX5 found in these studies was correlated with different stages of the peroxisomal matrix protein import cycle [33]. PEX5 molecules in stages 0 and 1 were protease-sensitive and localized in the cytosol and to peroxisome membranes, respectively. Stage 2 defined PEX5 molecules that were peroxisome membrane-associated and rendered 2 kDa shorter at their N-termini on protease treatment, indicating that these were embedded deep in the peroxisome membrane with about 2 kDa exposed to the cytosol. Finally, stage 3 PEX5 molecules were peroxisome-associated and completely protease-resistant, suggesting they might be in the lumen. ATP-limiting conditions, which inhibit matrix protein import and receptor recycling (as indicated later), enhance the amount of stage 3 molecules. These data are compatible with a model in which PEX5 begins in the cytosol (stage 0), docks at peroxisomes (stage 1), inserts into the peroxisome membrane (stage 2), and then translocates to the preoxisome matrix (stage 3), before its final recycling to the cytosol (stage 0), for another round of import. These intermediates in the shuttling of PEX5, as well as the ability of shuttling PEX5 to participate in more than one round of import, has also been documented beautifully in an *in vitro* system capable of importing proteins into the peroxisome matrix [24].

Do the receptor and coreceptor for the PTS2 pathway also exhibit the extended shuttle? Work on *S. cerevisiae* and *P. pastoris* shows this to be true for Pex7 and Pex20, respectively [17, 29]. The N-terminal 56 amino acids of Pex7 are necessary for its peroxisomal entry [14], and fusion of green fluorescent protein (GFP) or other tags at the C-terminus slows down or inhibits Pex7 recycling from peroxisomes [29]. Further details of Pex7 behavior during the import cycle are not as clear as they are for Pex5.

The auxiliary proteins (Pex18, Pex21, or Pex20) are required for PTS2 import, but different models have been proposed for their exact functions [17, 19]. In *S. cerevisiae*, Pex18 is suggested to stabilize the PTS2 cargo/Pex7

interactions by formation of a ternary complex that then docks at the peroxisome [19]. In *P. pastoris*, Pex7 and Pex20 can independently dock at the peroxisome membrane, but Pex20 is needed for the translocation of PTS2 cargo into the matrix [17]. Interestingly, the PTS2 pathway coreceptor in mammals, Pex5L, is also needed to translocate the Pex7/PTS2 cargo complex into peroxisomes [35].

The extended shuttle has been well documented for Pex20 from *P. pastoris* [17]. About two-thirds of Pex20-GFP is cytosolic in wild-type cells and the remaining one-third is peroxisomal and protease-protected, suggesting that it is in the matrix. In the absence of Pex14, a key constituent of the docking subcomplex, none of the Pex20-GFP associates with peroxisomes and all of it is cytosolic. Therefore, the docking subcomplex is necessary to translocate Pex20 into peroxiosomes. In *P. pastoris* cells lacking components of the RING subcomplex, all the Pex20-GFP is peroxisome-associated, suggesting that this subcomplex may aid the export of Pex20 from the peroxisome matrix. Finally, in the absence of any component of the receptor recycling machinery, both Pex5 and Pex20 cannot recycle from the peroxisomes to the cytosol. Under these conditions, one might expect them to accumulate on the peroxisome membrane just prior to relocation to the cytosol, but instead both Pex5 and Pex20 are degraded by a ubiquitin-proteaseome-dependent machinery responsible for a process we have dubbed RADAR [17].

The extended shuttle exhibited by the PTS1 and PTS2 receptors provides a satisfying explanation for the varying subcellular locations reported in the early studies on these proteins. The shuttle would still function as long as there is a dynamic exchange of receptors from the peroxisomes to the cytosol, irrespective of whether the predominant pool is in one or the other compartment.

The details of the receptor recycling process and the mechanism by which this is achieved are described elsewhere in this volume and will not be reiterated here. However, it is worth considering briefly how RADAR is activated and implemented.

## B. The Peroxisomal RADAR Pathway

In wild-type cells there appears to be little or no turnover of Pex5 or Pex20 as long as receptor recycling is active and functional [17, 36, 37]. It is possible that receptor recycling precedes RADAR, or that the latter machinery is inactive under these conditions. Upon a block in receptor recycling (and this is true of both Pex5 and Pex20), RADAR kicks in, and the receptors that fail to recycle from the peroxisome membrane are targeted for polyubiquitylation by UBCs other than Pex4 (most likely Ubc1, Ubc4, or Ubc5 in *S. cerevisiae*)

and degradation by the proteasome [17, 36, 37]. This RADAR pathway requires specific lysines, near the N-terminus of Pex5 and Pex20, which are the targets for polyubiquitylation [17, 38]. In *P. pastoris*, K22 in Pex5 and K19 in Pex20 are necessary for RADAR [17]. Mutations of one or other of these lysines abolish RADAR for that protein [17, 38]. Interestingly, the Pex18 protein from *S. cerevisiae* is constitutively degraded in a ubiquitin- and proteasome-dependent manner [39], which is characteristic of coreceptor engagement by the RADAR pathway [17, 36, 37]. We have found, in *P. pastoris*, that Pex20 mutants that are recycled to the cytosol inefficiently behave in exactly this manner, that is RADAR now clears the peroxisome surface of cargo-free receptors (Leon and Subramani, unpublished data). Notably, these PpPex20 mutants and endogenous ScPex18 are degraded by RADAR only if the import cycle is functional [39].

Mutation of these lysines (Pex5K22R or Pex20K19R) in *P. pastoris*, or of the corresponding lysine in *H. polymorpha* Pex5, does not affect peroxisome biogenesis, consistent with the lack of a significant role for RADAR under conditions when receptor recycling is functional [3, 38]. However, in the absence of receptor recycling (due to mutations in the receptors or in the receptor-recycling machinery), these proteins accumulate at the peroxisome membrane in their polyubiuquitylated forms, providing formal evidence for the requirement of Pex4, Pex22, Pex1, and Pex6 for receptor recycling from peroxisomes to the cytosol [3]. Thus, the effects of the RADAR pathway become apparent only when receptor recycling is compromised. This provides a clue to the physiological function of the RADAR pathway, which might be to clear, from the peroxisome membrane, cargo-free receptors that cannot be recycled by other mechanisms.

### C. PTS Receptor-Mediated Steps in the Matrix Protein Import Cycle

A complete understanding of the receptor shuttling pathway and mechanism requires an appreciation of the following steps (Figure 20.1):

1. *Cargo binding*: For the PTS1 pathway, most cargoes are bound by Pex5, via interactions of the TPR motifs on Pex5 with the C-terminal PTS1 peptide on the cargo [5, 40, 41]. In a few special cases, proteins such as *S. cerevisiae* acyl-CoA oxidase are imported into peroxisomes in a Pex5-dependent manner but this protein does not have a functional PTS1 [42]. Instead, the cargo protein interacts with a segment of Pex5 that is upstream of the canonical cargo-recognition site (i.e., the TPR repeats [42]).

Most PTS2 cargoes are recognized by Pex7 [12, 14], but may need proteins of the Pex20 family to stabilize this interaction, as reported for Pex18

of *S. cerevisiae* [19]. However, some cargoes, such as PpPex8 [43] and *H. polymorpha* amine oxidase [44], interact directly with Pex20. Therefore, the Pex20 family of proteins may be viewed as coreceptors for the PTS2 pathway. In plants and mammals, which do not have a Pex20-like protein, the Pex5L isoform serves as the coreceptor for the PTS2 pathway [15, 21]. Although most PTS2 proteins are recognized by their receptors or coreceptors interacting directly with the PTS2 sequence on the cargo, there are occasional exceptions to this rule, as noted above for the PTS1 cargoes. For example, *Y. lipolytica* thiolase interacts with Pex20 through a region outside the PTS2 [20].

2. *Receptor docking at the peroxisomal membrane*: Complexes formed between cargo and receptors/coreceptors interact with constituents of the docking subcomplex on the peroxisommal membrane [2]. Yeast Pex5, Pex7, and Pex20 are reported to interact independently with Pex13 and Pex14 of the docking subcomplex [3]. Pex5 and Pex20 also interact with Pex8 of this subcomplex. Although Pex17 is also a part of the docking subcomplex, none of the receptors or coreceptors appear to interact with this protein directly.

3. *Receptor/cargo translocation across the peroxisomal membrane*: It is quite likely that Pex14 and perhaps the whole docking subcomplex is necessary for the translocation of receptor/cargo complexes into the peroxisome matrix. A fraction of Pex5, Pex7, and Pex20 in yeasts is peroxisome-associated and protease-resistant, but in the absence of Pex14, Pex5 is not peroxisome-associated [43], Pex7 exhibits reduced binding to peroxisomes [45], and Pex20 is exclusively cytosolic [17]. The cotranslocation of receptor/cargo complexes is not surprising in view of the ability of peroxisomes to import folded and/or oligomeric proteins [3].

4. *Cargo release in the peroxisome matrix*: How this is achieved is not exactly clear. The environment in either the peroxisome matrix or specific proteins, such as Pex8, may be necessary, but it should be noted that the only intraperoxisomal peroxin, Pex8, is found only in yeasts, leaving open the question of how cargo release might occur in plants and mammals.

5. *Receptor export to the peroxisome membrane*: As predicted by Dammai and Subramani [27], both *cis*-acting sequences and *trans*-acting proteins are necessary for the export of receptors/coreceptors (presumably after cargo release) from the peroxisome matrix to the peroxisome membrane, from where they have to be recycled to the cytosol. The first 17 amino acids of human PEX5 are necessary for this step [46]. Similarly, the first 19 amino acids of Pex20 are required for its export [17]. Pex5 and Pex20 proteins share sequence similarities at their N-termini and these two proteins also exhibit similar dynamics and behavior during the matrix protein import cycle [3]. Hence, their mechanisms of export may be based on similar

principles. The nature of the sequences in Pex7 necessary for its export are not clear, but fusions to its C-terminus affect export efficiency [29].

Among the *trans*-acting proteins, components of the RING subcomplex, Pex2, Pex10, and Pex12, are necessary for the export of Pex20 [17]. In these mutants, Pex20 is peroxisome-associated but inaccessible to the RADAR machinery, suggesting that it is inside peroxisomes. In mammalian systems, the stage 2 state of Pex5 described earlier [47] may correspond to an intermediate that has been exported from the peroxisome matrix to the membrane, where only about 2 kDa of its N-terminus is exposed to the cytosol where it is clipped by protease. This model would be consistent with a role for the N-terminus in two distinct events—receptor export and recycling.

6. *Recycling of cargo-free receptors to the cytosol*: This probably requires monoubiquitylation of the N-termini of the Pex5 and Pex20 by Pex4 and the action of the two AAA ATPases, Pex1 and Pex6 [3, 25, 48]. The details of this process are reviewed separately in this volume.

7. *Clearance of the receptor from the peroxisome membrane when recycling is affected*: This process involves the RADAR pathway and has been described above.

## D. Energetics of Receptor Recycling and Cargo Import

Peroxisomal protein import is unusual in that the entry of both receptors and cargo is ATP independent [24, 25, 49]. However, the export of Pex5 from peroxisomes to the cytosol is ATP dependent [24, 25, 49]. Part of this requirement is that the ATP-dependent UBC, Pex4, probably monoubiquitylates the receptors, as a prelude to receptor export and/or recycling [3, 25, 48]. In addition, the ATPases, Pex1 and Pex6, act in a complex to hydrolyze ATP and facilitate the relocation of the receptors from the peroxisome to the cytosol [24, 25]. Finally, as described below, Hsp70, which also hydrolyzes ATP, is necessary for matrix protein import [50, 51]. It is still unclear exactly how many ATPs are consumed for the import of each molecule of cargo.

The receptor dynamics and the sites of the energy-requiring steps in peroxisomal matrix protein import are reminiscent of nuclear protein import, where cargo enters the nucleus in association with the NLS receptor, importin, but no GTP hydrolysis is required for this step per se [52]. The energy-requiring step in nuclear import is actually in the cytosol, where GTP hydrolysis on Ran-GTP releases the import receptor for another round of cargo binding and import. In peroxisomal protein import, ATP hydrolysis in the cytosol is also necessary for replenishing receptors in the cytosol for another round of cargo binding and import [24, 25, 49].

## III. Role of Hsp70 Family of Proteins in Peroxisomal Matrix Protein Import

The first evidence for the role of chaperones in peroxisomal matrix protein import came from a microinjection-based assay for the import of proteins into peroxisomes of mammalian cells [51]. This work showed an involvement of the constitutive (Hsp73), but not the inducible, heat-shock proteins (Hsps) of the 70-kDa family. Proteins of this family were associated with proteins being imported into the peroxisome matrix. Antibodies against the Hsp70 proteins inhibited peroxisomal import and the inhibition was reversed by the addition of exogenous Hsp70. This class of proteins was found associated with purified rat liver peroxisomes, and more peroxisome-associated Hsp70s were found during peroxisome proliferation. Interestingly, protease protection assays suggested that the association of Hsp73 terminated at the peroxisome membrane because Hsp73 remained protease-sensitive. However, when an unfolded protein, such as reduced, alkylated, and biotinylated human serum albumin conjugated to a peptide ending in the PTS1, SKL (bHSA-SKL), was microinjected into mammalian cells, Hsp70 associated with the cargo and was found inside the peroxisomes [53]. The rate of peroxisomal import of folded and unfolded bHSA-SKL was essentially the same [53]. The ability of chaperones to enter peroxisomes in association with unfolded proteins may circumvent the absence of chaperones in mammalian peroxisomes. The roles of intra- and extraperoxisomal Hsp70 in the import cycle may be quite distinct.

Proteins of the Hsp70 family function as ATPases, like the DnaK protein of *Escherichia coli* [54]. These proteins exist is two conformational states, one bound to ADP and with a higher affinity for protein/peptide binding, and the other bound to ATP and possessing a lower affinity for protein/peptide binding. The ATP hydrolysis rate, and hence the conversion of Hsp70 to the ADP-bound form, for this family of proteins is stimulated by members of the DnaJ, or Hsp40, family [55]. Thus, it was anticipated that members of this Hsp40 family would also be required for peroxisomal matrix protein import. In mammalian cells, this is indeed true as judged by the inhibition of import in a permeabilized mammalian cell system by antibodies specific to Hsp40 or Hsp70 [41]. Although the requirement of members of the Hsp70/Hsp40 family for peroxisomal matrix protein import was first demonstrated for PTS1-containing proteins [51], subsequent work in permeabilized mammalian cells revealed that this was also true for PTS2-containing proteins [56]. Thus, all or most peroxisomal matrix proteins need these chaperones for their import. Lending further support to this idea is the finding that among the mammalian organellar proteins that bound to *E. coli* Hsp70 were peroxisomal matrix proteins (the multifunctional enzyme and an isoform of 2,4-dienoyl-CoA reductase) [57].

The role of the Hsp70 family of proteins is not well studied in yeast systems that have been so instrumental in advancing our knowledge of the mechanism of peroxisome biogenesis. A member of the DnaJ family, Djp1p, has been reported to be involved specifically in the import of peroxisomal matrix proteins [58]. Cells lacking this protein were impaired only in peroxisomes and peroxisomal matrix protein import was affected to various extents. However, nuclear, endoplasmic reticulum (ER) and mitochondrial import were not affected.

Many members of the Hsp70 family also act in concert with nucleotide exchange factors that replace ADP with ATP (e.g., the BAG family of proteins in eukaryotes) [54]. No specific exchange factor has been identified to play a role in peroxisomal matrix protein import. However, there is indirect, suggestive evidence for the involvement of a plant nucleotide exchange factor in the targeting of a peroxisomal membrane protein, cottonseed ascorbate peroxidase (APX), to the ER, from where it is proposed to be sorted to peroxisomes [59].

In plants, there are several reports of the existence of Hsp70 family proteins on or in peroxisomes, as well as of their requirement for peroxisome biogenesis. An Hsp70 homologue from *Citrullus vulgaris* was shown to be targeted to both chloroplasts and peroxisomes by initiation of protein translation at two different methionines [60]. The peroxisomal variant has a PTS2 sequence. A DnaJ (Hsp40) homologue from *Cucumis sativus* was found to be attached to the glyoxysomal membrane, in association with a cytosolic Hsp70 [61]. Plant peroxisomal protein import *in vitro* is enhanced by the presence of chaperones [50]. Peroxisomes isolated from heat-shocked pumpkin seedlings showed more import than those from untreated seedlings. Additionally, antibodies to wheat germ Hsp70 (and *E. co*li Hsp90) inhibited peroxisomal import of isocitrate lyase. Finally, immunoprecipitates of Hsp70 contained peroxisomal matrix proteins.

The posttranslational import of a plant peroxisomal membrane protein, APX, into purified ER membranes (from where they are proposed to sort to the peroxisomes) *in vitro*, was also impaired by the immunodepletion of Hsp70, AtJ2 (a DnaJ homologue), and AtE1 (an *E. coli* GrpE homologue) [59].

Small heat-shock proteins (sHsps; 16–42 kDa) with PTS1- and PTS2-targeting sequences have been described in plants [62]. Although the function of these proteins is unclear in plants, in other systems they prevent protein aggregation.

In other organelles that transport unfolded proteins across their membranes (ER, mitochondria, and chloroplasts), the requirement for Hsp70/40 proteins is not surprising because these chaperones help to maintain the newly synthesized proteins in the unfolded state prior to translocation

across the membranes. However, the peroxisomal translocon is unusual in that folded, oligomeric and cofactor-bound proteins can be translocated across the peroxisomal membrane [3]. In view of this, the exact role of Hsp70 remained an enigma but three models have been proposed [63].

In the first model, the chaperones might stabilize the exposed C-terminal PTS1 sequence during thermal fluctuations that unfold the C-terminal end of the protein, without unfolding the rest of the protein. This would predict that the presence of chaperones might aid the interactions between the PTS receptor and its cargo. The second model suggests that Hsp70 proteins might help to facilitate assembly of the cargo/receptor complexes with docking proteins on the surface of peroxisomes. The final model is one where Hsp70/Hsp40 act in a manner analogous to clathrin-uncoating ATPase, to disassemble protein complexes on the peroxisome membrane. In the light of current knowledge regarding the extended shuttling and recycling of PTS receptors, an obvious possibility is a late requirement for Hsp70 in the receptor-release step at the peroxisome membrane.

The binding between the purified PTS1 receptor domain that binds cargo and a PTS1 peptide was studied by fluorescence anisotropy [64]. Specific binding of the receptor to the PTS1 peptide was observed, but no effect was seen on addition of Hsp70, Hsp70 and ATP, or Hsp70 and ADP. However, in these experiments, no Hsp40 was added. In a conflicting report, it was shown that mammalian Hsp70 interacts with the cargo-binding (TPR) domain of PEX5, and that Hsp70 and ATP synergistically enhance the binding of PEX5 to the PTS1 of acyl-CoA oxidase, a peroxisomal matrix protein [65]. However, the energy dependence of cargo binding is not consistent with other reports that cargo binding and import into peroxisomes do not require ATP [24].

Therefore, at present it is still unclear exactly how and when Hsp70/Hsp40 proteins act during peroxisomal matrix protein import.

## Acknowledgments

This work was supported by grants DK41737 and GM59844 to SS.

## References

1. Heiland, I., and Erdmann, R. (2005). Biogenesis of peroxisomes. Topogenesis of the peroxisomal membrane and matrix proteins. *FEBS J* 272:2362–2372.
2. Rayapuram, N., and Subramani, S. (2006). The importomer—a peroxisomal membrane complex involved in protein translocation into the peroxisome matrix. *Biochim Biophys Acta* 1763(12):1613–1619.

3. Leon, S., Goodman, J.M., and Subramani, S. (2006). Uniqueness of the mechanism of protein import into the peroxisome matrix: Transport of folded, co-factor-bound and oligomeric proteins by shuttling receptors. *Biochim Biophys Acta* 1763(12):1552–1564.
4. Dodt, G., Braverman, N., Wong, C., Moser, A., Moser, H.W., Watkins, P., Valle, D., and Gould, S.J. (1995). Mutations in the PTS1 receptor gene, *PXR1*, define complementation group 2 of the peroxisome biogenesis disorders. *Nat Genet* 9:115–125.
5. McCollum, D., Monosov, E., and Subramani, S. (1993). The *pas8* mutant of *Pichia pastoris* exhibits the peroxisomal protein import deficiencies of Zellweger syndrome cells—the PAS8 protein binds to the COOH-terminal tripeptide peroxisomal targeting signal, and is a member of the TPR protein family. *J Cell Biol* 121:761–774.
6. Terlecky, S.R., Nuttley, W.M., McCollum, D., Sock, E., and Subramani, S. (1995). The *Pichia pastoris* peroxisomal protein PAS8p is the receptor for the C-terminal tripeptide peroxisomal targeting signal. *EMBO J* 14:3627–3634.
7. Wiemer, E.A., Nuttley, W.M., Bertolaet, B.L., Li, X., Francke, U., Wheelock, M.J., Anne, U.K., Johnson, K.R., and Subramani, S. (1995). Human peroxisomal targeting signal-1 receptor restores peroxisomal protein import in cells from patients with fatal peroxisomal disorders. *J Cell Biol* 130:51–65.
8. Braverman, N., Steel, G., Obie, C., Moser, A., Moser, H., Gould, S.J., and Valle, D. (1997). Human *PEX7* encodes the peroxisomal PTS2 receptor and is responsible for rhizomelic chondrodysplasia punctata. *Nat Genet* 15:369–376.
9. Elgersma, Y., Elgersma-Hooisma, M., Wenzel, T., McCaffery, J.M., Farquhar, M.G., and Subramani, S. (1998). A mobile PTS2 receptor for peroxisomal protein import in *Pichia pastoris*. *J Cell Biol* 140:807–820.
10. Marzioch, M., Erdmann, R., Veenhuis, M., and Kunau, W.H. (1994). *PAS7* encodes a novel yeast member of the WD-40 protein family essential for import of 3-oxoacyl-CoA thiolase, a PTS2-containing protein, into peroxisomes. *EMBO J* 13:4908–4918.
11. Purdue, P.E., Zhang, J.W., Skoneczny, M., and Lazarow, P.B. (1997). Rhizomelic chondrodysplasia punctata is caused by deficiency of human PEX7, a homologue of the yeast PTS2 receptor. *Nat Genet* 15:381–384.
12. Rehling, P., Marzioch, M., Niesen, F., Wittke, E., Veenhuis, M., and Kunau, W.H. (1996). The import receptor for the peroxisomal targeting signal 2 (PTS2) in *Saccharomyces cerevisiae* is encoded by the *PAS7* gene. *EMBO J* 15:2901–2913.
13. Zhang, J.W., and Lazarow, P.B. (1995). *PEB1* (*PAS7)* in *Saccharomyces cerevisiae* encodes a hydrophilic, intra-peroxisomal protein that is a member of the WD repeat family and is essential for the import of thiolase into peroxisomes. *J Cell Biol* 129:65–80.
14. Zhang, J.W., and Lazarow, P.B. (1996). Peb1p (Pas7p) is an intraperoxisomal receptor for the $NH_2$-terminal, type 2, peroxisomal targeting sequence of thiolase: Peb1p itself is targeted to peroxisomes by an NH2-terminal peptide. *J Cell Biol* 132:325–334.
15. Woodward, A.W., and Bartel, B. (2005). The *Arabidopsis* peroxisomal targeting signal type 2 receptor PEX7 Is necessary for peroxisome function and dependent on PEX5. *Mol Biol Cell* 16:573–583.
16. Einwächter, H., Sowinski, S., Kunau, W.H., and Schliebs, W. (2001). *Yarrowia lipolytica* Pex20p, *Saccharomyces cerevisiae* Pex18p/Pex21p and mammalian Pex5pL fulfil a common function in the early steps of the peroxisomal PTS2 import pathway. *EMBO Rep* 2:1035–1039.
17. Leon, S., Zhang, L., Mcdonald, W.H., Yates, J., III, Cregg, J.M., and Subramani, S. (2006). Dynamics of the peroxisomal import cycle of PpPex20p: ubiquitin-dependent localization and regulation. *J Cell Biol* 172:67–78.

18. Purdue, P.E., Yang, X., and Lazarow, P.B. (1998). Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. *J Cell Biol* 143:1859–1869.
19. Stein, K., Schell-Steven, A., Erdmann, R., and Rottensteiner, H. (2002). Interactions of Pex7p and Pex18p/Pex21p with the peroxisomal docking machinery: implications for the first steps in PTS2 protein import. *Mol Cell Biol* 22:6056–6069.
20. Titorenko, V.I., Smith, J.J., Szilard, R.K., and Rachubinski, R.A. (1998). Pex20p of the yeast *Yarrowia lipolytica* is required for the oligomerization of thiolase in the cytosol and for its targeting to the peroxisome. *J Cell Biol* 142:403–420.
21. Braverman, N., Dodt, G., Gould, S.J., and Valle, D. (1998). An isoform of Pex5p, the human PTS1 receptor, is required for the import of PTS2 proteins into peroxisomes. *Hum Mol Genet* 7:1195–1205.
22. Agne, B., Meindl, N.M., Niederhoff, K., Einwachter, H., Rehling, P., Sickmann, A., Meyer, H.E., Girzalsky, W., and Kunau, W.H. (2003). Pex8p: an intraperoxisomal organizer of the peroxisomal import machinery. *Mol Cell* 11:635–646.
23. Hazra, P.P., Suriapranata, I., Snyder, W.B., and Subramani, S. (2002). Peroxisome remnants in *pex3Δ* cells and the requirement of Pex3p for interactions between the peroxisomal docking and translocation subcomplexes. *Traffic* 3:560–574.
24. Miyata, N., and Fujiki, Y. (2005). Shuttling mechanism of peroxisome targeting signal type 1 receptor Pex5: ATP-independent import and ATP-dependent export. *Mol Cell Biol* 25:10822–10832.
25. Platta, H.W., Grunau, S., Rosenkranz, K., Girzalsky, W., and Erdmann, R. (2005). Functional role of the AAA peroxins in dislocation of the cycling PTS1 receptor back to the cytosol. *Nat Cell Biol* 7:817–822.
26. Collins, C.S., Kalish, J.E., Morrell, J.C., Mccaffery, J.M., and Gould, S.J. (2000). The peroxisome biogenesis factors Pex4p, Pex22p, Pex1p, and Pex6p act in the terminal steps of peroxisomal matrix protein import. *Mol Cell Biol* 20:7516–7526.
27. Dammai, V., and Subramani, S. (2001). The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* 105:187–196.
28. Kunau, W.H. (2001). Peroxisomes: the extended shuttle to the peroxisome matrix. *Curr Biol* 11:R659–R662.
29. Nair, D.M., Purdue, P.E., and Lazarow, P.B. (2004). Pex7p translocates in and out of peroxisomes in *Saccharomyces cerevisiae*. *J Cell Biol* 167:599–604.
30. Rachubinski, R.A., and Subramani, S. (1995). How proteins penetrate peroxisomes. *Cell* 83:525–528.
31. Dodt, G., and Gould, S.J. (1996). Multiple *PEX* genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor. *J Cell Biol* 135:1763–1774.
32. Van Der Klei, I.J., Hilbrands, R.E., Kiel, J.A., Rasmussen, S.W., Cregg, J.M., and Veenhuis, M. (1998). The ubiquitin-conjugating enzyme Pex4p of *Hansenula polymorpha* is required for efficient functioning of the PTS1 import machinery. *EMBO J* 17:3608–3618.
33. Gouveia, A.M., Guimaraes, C.P., Oliveira, M.E., Sa-Miranda, C., and Azevedo, J.E. (2003). Insertion of Pex5p into the peroxisomal membrane is cargo protein-dependent. *J Biol Chem* 278:4389–4392.
34. Gouveia, A.M., Reguenga, C., Oliveira, M.E., Sa-Miranda, C., and Azevedo, J.E. (2000). Characterization of peroxisomal Pex5p from rat liver. Pex5p in the Pex5p-Pex14p membrane complex is a transmembrane protein. *J Biol Chem* 275:32444–32451.

35. Otera, H., Harano, T., Honsho, M., Ghaedi, K., Mukai, S., Tanaka, A., Kawai, A., Shimizu, N., and Fujiki, Y. (2000). The mammalian peroxin Pex5pL, the longer isoform of the mobile peroxisome targeting signal (PTS) type 1 transporter, translocates the Pex7p. PTS2 protein complex into peroxisomes via its initial docking site, Pex14p. *J Biol Chem* 275:21703–21714.
36. Kiel, J.A., Emmrich, K., Meyer, H.E., and Kunau, W.H. (2005). Ubiquitination of the peroxisomal targeting signal type 1 receptor, Pex5p, suggests the presence of a quality control mechanism during peroxisomal matrix protein import. *J Biol Chem* 280:1921–1930.
37. Platta, H.W., Girzalsky, W., and Erdmann, R. (2004). Ubiquitination of the peroxisomal import receptor Pex5p. *Biochem J* 384:37–45.
38. Kiel, J.A., Otzen, M., Veenhuis, M., and Van Der Klei, I.J. (2005). Obstruction of polyubiquitination affects PTS1 peroxisomal matrix protein import. *Biochim Biophys Acta* 1745:176–186.
39. Purdue, P.E., and Lazarow, P.B. (2001). Pex18p is constitutively degraded during peroxisome biogenesis. *J Biol Chem* 276:47684–47689.
40. Maynard, E.L., Gatto, G.J., Jr., and Berg, J.M. (2004). Pex5p binding affinities for canonical and noncanonical PTS1 peptides. *Proteins* 55:856–861.
41. Terlecky, S.R., Wiemer, E.A., Nuttley, W.M., Walton, P.A., and Subramani, S. (1996). Signals, receptors, and cytosolic factors involved in peroxisomal protein import. *Ann N Y Acad Sci* 804:11–20.
42. Klein, A.T., Van Den Berg, M., Bottger, G., Tabak, H.F., and Distel, B. (2002). *Saccharomyces cerevisiae* acyl-CoA oxidase follows a novel, non-PTS1, import pathway into peroxisomes that is dependent on Pex5p. *J Biol Chem* 277:25011–25019.
43. Zhang, L., Leon, S., and Subramani, S. (2006). Two independent pathways traffic the intraperoxisomal peroxin PpPex8p into peroxisomes: mechanism and evolutionary implications. *Mol Biol Cell* 17:690–699.
44. Otzen, M., Wang, D., Lunenborg, M.G., and Van Der Klei, I.J. (2005). *Hansenula polymorpha* Pex20p is an oligomer that binds the peroxisomal targeting signal 2 (PTS2). *J Cell Sci* 118:3409–3418.
45. Girzalsky, W., Rehling, P., Stein, K., Kipper, J., Blank, L., Kunau, W.H., and Erdmann, R. (1999). Involvement of Pex13p in Pex14p localization and peroxisomal targeting signal 2-dependent protein import into peroxisomes. *J Cell Biol* 144:1151–1162.
46. Costa-Rodrigues, J., Carvalho, A.F., Gouveia, A.M., Fransen, M., Sa-Miranda, C., and Azevedo, J.E. (2004). The N terminus of the peroxisomal cycling receptor, Pex5p, is required for redirecting the peroxisome-associated peroxin back to the cytosol. *J Biol Chem* 279:46573–46579.
47. Gouveia, A.M., Guimaraes, C.P., Oliveira, M.E., Reguenga, C., Sa-Miranda, C., and Azevedo, J.E. (2003). Characterization of the peroxisomal cycling receptor, Pex5p, using a cell-free *in vitro* import system. *J Biol Chem* 278:226–230.
48. Kragt, A., Voorn-Brouwer, T., Van Den Berg, M., and Distel, B. (2005). The *Saccharomyces cerevisiae* peroxisomal import receptor Pex5p is monoubiquitinated in wild type cells. *J Biol Chem* 280:7867–7874.
49. Oliveira, M.E., Gouveia, A.M., Pinto, R.A., Sa-Miranda, C., and Azevedo, J.E. (2003). The energetics of Pex5p-mediated peroxisomal protein import. *J Biol Chem* 278:39483–39488.
50. Crookes, W.J., and Olsen, L.J. (1998). The effects of chaperones and the influence of protein assembly on peroxisomal protein import. *J Biol Chem* 273:17236–17242.

51. Walton, P.A., Wendland, M., Subramani, S., Rachubinski, R.A., and Welch, W.J. (1994). Involvement of 70-kD heat-shock proteins in peroxisomal import. *J Cell Biol* 125:1037–1046.
52. Gorlich, D. (1997). Nuclear protein import. *Curr Opin Cell Biol* 9:412–419.
53. Brocard, C.B., Jedeszko, C., Song, H.C., Terlecky, S.R., and Walton, P.A. (2003). Protein structure and import into the peroxisomal matrix. *Traffic* 4:74–82.
54. Bukau, B., Weissman, J., and Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell* 125:443–451.
55. Craig, E.A., Huang, P., Aron, R., and Andrew, A. (2006). The diverse roles of J-proteins, the obligate Hsp70 co-chaperone. *Rev Physiol Biochem Pharmacol* 156:1–21.
56. Legakis, J.E., and Terlecky, S.R. (2001). PTS2 protein import into mammalian peroxisomes. *Traffic* 2:252–260.
57. Naylor, D.J., Hoogenraad, N.J., and Hoj, P.B. (1999). Characterisation of several Hsp70 interacting proteins from mammalian organelles. *Biochim Biophys Acta* 1431:443–450.
58. Hettema, E.H., Ruigrok, C.C., Koerkamp, M.G., Van Den Berg, M., Tabak, H.F., Distel, B., and Braakman, I. (1998). The cytosolic DnaJ-like protein Djp1p is involved specifically in peroxisomal protein import. *J Cell Biol* 142:421–434.
59. Mullen, R.T., Lisenbee, C.S., Miernyk, J.A., and Trelease, R.N. (1999). Peroxisomal membrane ascorbate peroxidase is sorted to a membranous network that resembles a subdomain of the endoplasmic reticulum. *Plant Cell* 11:2167–2185.
60. Wimmer, B., Lottspeich, F., Van Der Klei, I., Veenhuis, M., and Gietl, C. (1997). The glyoxysomal and plastid molecular chaperones (70-kDa heat shock protein) of watermelon cotyledons are encoded by a single gene. *Proc Natl Acad Sci USA* 94:13624–13629.
61. Diefenbach, J., and Kindl, H. (2000). The membrane-bound DnaJ protein located at the cytosolic site of glyoxysomes specifically binds the cytosolic isoform 1 of Hsp70 but not other Hsp70 species. *Eur J Biochem* 267:746–754.
62. Ma, C., Haslbeck, M., Babujee, L., Jahn, O., and Reumann, S. (2006). Identification and characterization of a stress-inducible and a constitutive small heat-shock protein targeted to the matrix of plant peroxisomes. *Plant Physiol* 141:47–60.
63. Subramani, S. (1998). Components involved in peroxisome import, biogenesis, proliferation, turnover, and movement. *Physiol Rev* 78:171–188.
64. Harper, C.C., Berg, J.M., and Gould, S.J. (2003). PEX5 binds the PTS1 independently of Hsp70 and the peroxin PEX12. *J Biol Chem* 278:7897–7901.
65. Harano, T., Nose, S., Uezu, R., Shimizu, N., and Fujiki, Y. (2001). Hsp70 regulates the interaction between the peroxisome targeting signal type 1 (PTS1)-receptor Pex5p and PTS1. *Biochem J* 357:157–165.

# 21

# *Function of the Ubiquitin-Conjugating Enzyme Pex4p and the AAA Peroxin Complex Pex1p/Pex6p in Peroxisomal Matrix Protein Transport*

HARALD W. PLATTA • SVEN THOMS •
WOLF-H. KUNAU • RALF ERDMANN

*Abteilung für Systembiochemie*
*Medizinische Fakultät der Ruhr-Universität Bochum*
*D-44780 Bochum, Germany*

## I. Introduction

Peroxisomes are cellular organelles present in all eukaryotic cells. They compartmentalize many enzymes involved in lipid metabolism and defense against oxidative stress [1, 2]. The importance of peroxisome-specific metabolic processes is reflected by human disorders associated with peroxisomal defects [3]. These genetically determined disorders can be categorized as (1) disorders resulting from a defect in a single peroxisomal metabolic enzyme like the X-linked adrenoleukodystrophy or (2) disorders that result from a deficiency in the biogenesis of the peroxisome, referred to as peroxisome biogenesis disorders (PBDs) with the Zellweger syndrome being the best known example. The incidence of PBDs is rather low; however, there is no cure to date for these fatal diseases that normally lead to the death of the affected children within their first year of life.



ISSN NO: 1874-6047
DOI: 10.1016/S1874-6047(07)25021-8

Peroxisomes were first described by Rhodin [4] as ''microbodies'' in mouse kidney. Subsequently, DeDuve [5] isolated an organelle fraction containing catalase and $H_2O_2$-producing oxidases which he named peroxisomes. During the following two decades, advances in our understanding of the biochemistry of peroxisomes and their important role in β-oxidation (all organisms), glyoxylate cycle (plants and fungi), photorespiration (plants), ether lipid and cholesterol biosynthesis (mammals), and glycolysis (trypanosomatids) were achieved [3, 6–8]. This pronounced variability in enzyme content and thus metabolic function marked peroxisomes as multipurpose organelles. Additionally, they vary in morphology, size, and number depending on the species or cell type. The study of peroxisomal biogenesis and protein import was hampered for a long time by the low abundance of peroxisomes in many tissues and their great fragility. This changed when it became clear that peroxisome proliferation could be induced by manipulation of the carbon source in baker's yeast [9]. When offered a fatty acid as the sole carbon source, peroxisomes are required for growth because they are the exclusive site for fatty acid degradation in yeast. This allowed the screening and identification of mutants in peroxisome function with genetic approaches [10, 11]. Mutants affected in peroxisome biogenesis are phenotypically characterized by the mislocalization of peroxisomal matrix proteins and referred to as *pex* mutants, which identifies the corresponding gene as a *PEX* gene and the protein it encodes as a peroxin [12]. To date 32 peroxins are known [13]. They are involved in the three key stages of peroxisomal development: (1) formation of the peroxisomal membrane (import of peroxisomal membrane proteins), (2) compartmentalization of peroxisomal matrix enzymes (import of peroxisomal matrix proteins), and (3) peroxisome proliferation.

This chapter will focus on the enzymatically catalyzed mechanisms underlying transport of matrix proteins across the peroxisomal membrane into the lumen of the organelle, a process that involves most of the known peroxins.

## II. Peroxisomal Matrix Protein Import

Peroxisomes are surrounded by a single membrane and lack DNA. This means that their proteins are encoded in the nucleus, and the matrix proteins and most of the membrane proteins are synthesized on free ribosomes and imported posttranslationally (reviewed in [14]), while accumulating data indicate that the peroxisomal membranes and a subset of the peroxisomal membrane proteins might derive from the endoplasmic reticulum (ER) (reviewed in [13]). After formation of the protein import machinery at the

peroxisomal membrane, the matrix proteins are recognized by dynamic receptors in the cytosol and directed to the peroxisome. According to the model of shuttling receptors, the receptor/cargo complex reaches the luminal side of the peroxisome, where the complex is disassembled in order to release the cargo and the receptor is returned to the cytosol.

### A. Import of Folded and Oligomeric Proteins Across the Peroxisomal Membrane

Another characteristic feature of peroxisomes is the fact that they can import fully folded and even oligomeric protein complexes. This distinguishes them from the well-established translocons of mitochondria, chloroplasts, and the ER, which import unfolded polypeptides only. One example is catalase that can cross the peroxisomal membrane as a tetramer [15]. The abundant matrix protein alcohol oxidase had been shown to be imported in *Candida bondinii* as fully folded monomers that form homo-octameric complexes soon after import [16]. Another example is acyl-CoA oxidase, which was shown to be imported as a pentameric complex in *Yarrowia lipolytica* [17].

### B. Sequential Model for PTS-Receptor Cycle

The import of matrix enzymes is accomplished by receptors that shuttle between the cytosol and the peroxisomal compartment. In the following sections, we will briefly summarize the basic steps of this receptor cycle (Figure 21.1).

#### *1. Receptor-Substrate Binding*

In principle, proteins destined for import into peroxisomes are targeted via two pathways that rely on two conserved peroxisomal targeting signals (PTSs). The majority of peroxisomal matrix proteins possess a PTS1 at the very C-terminus consisting of the tripeptide SKL sequence or species-specific variants [18]. Pex5p interacts with the signal via six tetratricopeptide repeats (TPRs) within its C-terminal half [19]. The PTS2 is found near the N-terminus of only a few matrix proteins and has the consensus sequence (R/K)/(L/V/I)X5(H(Q)(L/A) [20]. PTS2-containing proteins are recognized by the WD40 protein Pex7p [21, 22]. Pex7p cooperates with the coreceptors Pex18p and Pex21p in *Saccharomyces cerevisiae* [23] or its orthologues Pex20p-like proteins in different yeast and fungi species [24–27] in order to form the receptor/cargo complex. In mammalian cells, a short and a long isoform of Pex5p have been identified, termed Pex5S and Pex5L [28, 29]. These proteins differ only in a short insertion, which

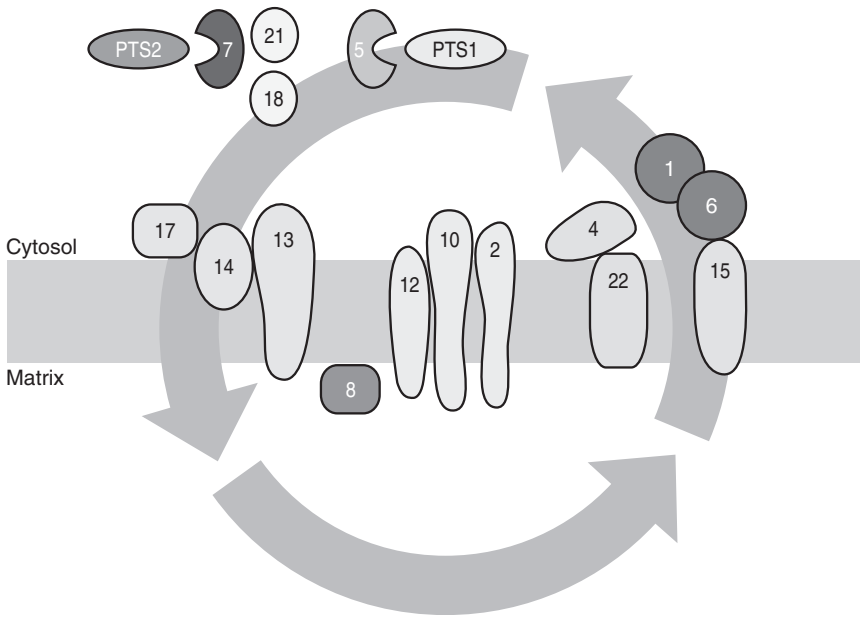



FIG. 21.1. PTS receptor cycle during peroxisomal matrix protein import. The peroxisomal matrix protein import receptor Pex5p recognizes cargo proteins harboring the PTS1. Cargo proteins containing the PTS2 sequence bind to the PTS2 receptor Pex7p, which is associated with the coreceptors Pex18p and Pex21p in *S. cerevisiae*. The receptor/cargo complexes reach the peroxisomal membrane and associate with the docking complex, consisting of the integral membrane protein Pex13p as well as the peripheral Pex14p and Pex17p. The PTS receptors reach the peroxisomal lumen, release their cargo, and are dislocated back to the cytosol. The RING-finger motif-containing peroxins Pex2p, Pex10p, and Pex12p form a complex that is connected to the docking factors via Pex8p. The RING-finger proteins are supposed to be involved in the process of receptor dislocation to the cytosol. Pex10p is linked to the ubiquitin-conjugating enzyme Pex4p, which is anchored to the peroxisomal membrane by Pex22p. The AAA proteins Pex1p and Pex6p are involved in PTS receptor release. They interact with each other and are attached to the peroxisomal membrane via Pex15p.

has been demonstrated to bind Pex7p. Furthermore, it has been shown that PTS1 and PTS2 pathways are also coupled in plants [30, 31]. Consequently, while in yeast und fungi both PTS pathways share the same membrane-bound protein import machinery, they converge in higher eukaryotes already at the level of Pex5L [32].

### 2. *Docking and Membrane Association of the Receptor/Cargo Complex*

The docking complex of the peroxisomal import machinery for matrix proteins comprises the three peroxins Pex13p, Pex14p, and Pex17p. Pex17p is a peripheral membrane protein that associates to peroxisomes via

Pex14p [33]. Pex13p is an integral membrane protein that contains an Src homology (SH3) domain that directly binds Pex5p as well as Pex14p, whereas the N-terminal region is required for interaction with Pex7p [34]. Pex5p and Pex20p contain Wxxx(F/Y) repeats that have been demonstrated to be needed for interaction with the docking factors [27, 35–37]. The amount of Pex5p, Pex7p, and Pex20p that can be found at peroxisomal membrane remnants is dramatically reduced in *pex14*Δ mutant cells in comparison to *pex13*Δ, *pex12*Δ or wild-type membranes [29]. Additionally cargo-loaded Pex5p exhibits a higher binding affinity to Pex14p than Pex13p [38]. Thus, the interaction to Pex14p is believed to be the entrance for Pex5p to the complex network of protein–protein interactions at the peroxisomal membrane.

### *3. Translocation and Cargo Release*

The exact mechanism underlying translocation and the components of the translocon have not yet been identified. The possibility is discussed that components of the docking subcomplex themselves are part of the translocon [39]. The multiple binding sites for Pex5p at the peroxisomal membrane might reflect the existence of an import cascade where cargo-loaded receptor interacts with different components of the import machinery [40].

One interesting fact about Pex5p is that it changes its membrane topology during protein import cascade as it behaves like an integral membrane protein in concert with Pex14p [41]. Finally, Pex5p enters the lumen of the peroxisome [42], although it is still a matter of debate if the whole receptor/cargo complex (extended shuttle hypothesis) or just a part of Pex5p (simple shuttle hypothesis) does see the matrix during cargo release [43]. Pex7p has been demonstrated to behave like a cycling receptor as well [44]. A model was postulated drawing parallels to pore forming toxins and suggesting that a population of Pex5p itself forms the dynamic import pore via protein–lipid interactions, thereby opening the membrane dynamically for a cargo-loaded second Pex5p species [39].

The exact mechanism of cargo release inside the peroxisome is not understood. The functional role of the intraperoxisomal Pex8p, which contains both the PTS1- as well as the PTS2-sequence that may function in disassembling the receptor/cargo complex, is discussed controversly [45, 46].

### *4. Receptor Release to the Cytosol*

Following to cargo liberation, Pex5p, Pex7p, and Pex20p are exported to the cytosol again. This process may need the peroxins Pex2p, Pex10p, and Pex12p, which are RING (*really interesting new gene*)-finger-containing proteins. This model is based on the observation that Pex5p and Pex20p

accumulate in the peroxisomal lumen in RING-finger mutants [27, 47–49]. This RING-finger complex itself is linked to the docking complex via Pex8p in *S. cerevisiae* [47].

Pex4p is a ubiquitin-conjugating enzyme [50], which is anchored to the peroxisomal membrane via Pex22p [51–52]. An epistasis analysis placed Pex4p and Pex22p at the end of the Pex5p receptor cycle together with the AAA peroxins Pex1p and Pex6p [53]. The AAA (*ATPases associated with various cellular activities*) peroxins form a complex in the cytosol [54] and are linked to the peroxisomal membrane via Pex15p in *S. cerevisiae* [55] or the orthologue Pex26p in mammalian cells [56]. Deletion of these components leads to an accumulation of Pex5p at the peroxisomal membrane and impairs Pex5p export *in vitro* in cells of *S. cerevisiae* [57] and human patients [58]. After remerging to the cytosol, the receptors are available for further rounds of protein import.

While the import of Pex5p appears to be based on binding affinity-mediated protein–protein or, according to the ''transient pore model,'' additional protein–lipid interactions [39], the export relies on three peroxisome specific enzymatic activities: ubiquitin-conjugating enzymes (Pex4p), ubiquitin-protein isopeptide ligases (putatively the RING-finger peroxins), and ATPases (the AAA peroxins Pex1p and Pex6p).

## III. Overview: Enzymes Involved in Ubiquitination

### A. Enzymatic Cascade for Protein Modification

Ubiquitination takes place as an enzymological cascade requiring the sequential activity of three types of enzymes. The ubiquitin-activating enzyme (Uba or E1) hydrolyzes ATP and forms a high-energy thioester bond between the cysteine of its active site and the C-terminal Gly76 of ubiquitin, thereby activating the COOH-terminus of ubiquitin for nucleophilic attack. Activated ubiquitin is then passed on to a ubiquitin-conjugating enzyme (Ubc or E2) by transthiolation. Finally, ubiquitin is covalently attached to the ε-amino group of a target lysine residue within the substrate protein through an isopeptide bond. This last reaction is mediated by a ubiquitin-protein isopeptide ligase (E3), which provides substrate specificity for this reaction, as it binds directly to the target protein. While homologous to E6-AP C-terminus (HECT)-E3 enzymes can form a thioester with ubiquitin themselves before substrate attachment, RING- and U-box E3s do not seem to interact with ubiquitin and primarily serve as platforms to coordinate E2 and substrate [59]. The organization of this reaction sequence is hierarchical: Most species contain only one single

E1 enzyme, while in yeast 11 and in humans about 100 E2 enzymes are known. The number of ubiquitin-ligases is still growing and is anticipated to be around 50 in yeast and about 1000 in human cells [60].

The types of ubiquitin modifications that can form are diverse. In the simplest form, one single ubiquitin molecule is attached, which is defined as monoubiquitination. Alternatively, several lysine residues can be charged with single ubiquitin molecules, giving rise to multiple monoubiquitination (multiubiquitination). Since ubiquitin contains seven lysine residues itself, ubiquitin molecules can become substrates for ubiquitin-ligation themselves and are able to form different types of chains in a process known as polyubiquitination.

### B. Downstream Components of Ubiquitin-Based Protein-Targeting Systems

Ubiquitin is an inducible and reversible signal. The different kinds of ubiquitin modifications function as signaling-dependent devices for regulating cellular functions [61]. Modification of target proteins with polyubiquitin chains linked via K48 are the best characterized to date. This modification serves as the classical degradation signal for cytosolic and organellar proteins as K48-branched chains of at least four ubiquitin molecules are recognized by the 26S proteasome. K63-linked chains and monoubiquitination are involved in internalization during endocytosis of plasma membrane receptors. Monoubiquitin is further associated with membrane protein uptake in late endosome vesicles for delivery to the lysosome or vacuole [62]. Especially (multiple) monoubiquitination events are linked to a wide range of nonproteolytical processes by serving as targeting signals [63]. Examples include p53 or Rad18 that alternatively need to be monoubiquitinated for export from the nucleus or they are polyubiquitinated for proteasomal degradation [61, 64, 65]. Monoubiquitination of histones is required for mitotic cell growth and meiosis and regulation of eukaryotic transcription activation [66]. DNA repair (nucleotide excision repair) is linked to transient posttranslational modifications of PCNA, which is marked by monoubiquitination, a K63-linked polyubiquitin chain or by the ubiquitin-like protein SUMO at the same lysine residue [67].

The different forms of modification are recognized by proteins containing specialized ubiquitin-binding domains (UBDs) that noncovalently bind to ubiquitin and thus function as downstream components of the ubiquitin-targeting system. They all have different ubiquitin-binding affinities, depending on the nature of the ubiquitin modification [60]. The conjugation of ubiquitin to its target protein is a reversible process. Removal of ubiquitin is mediated by deubiquitinating enzymes (DUBs). Besides recycling of free

ubiquitin molecules, DUBs function in stabilization of proteins or quenching of ubiquitin-based targeting signals [68].

## IV. The Ubiquitin-Conjugating Enzyme Pex4p in Peroxisome Biogenesis

### A. Ubiquitin-Conjugating Enzymes of the Ubc4p Family Involved in PTS Receptor Regulation

Ubiquitin-conjugating enzymes play a central role in the process of ubiquitination. They function to bridge the first, nonspecific step of ubiquitin activation by E1 with the transfer of activated ubiquitin to target-proteins by substrate-specific E3 enzymes.

Peroxisomal biogenesis has been implicated with ubiquitination ever since the discovery of the ubiquitin-conjugating enzyme Pex4p in 1992 [50], which had also been classified as Ubc10p [59]. Studies identified the Ubc1p/Ubc4p/Ubc5p familiy to be involved in peroxisomal processes in *S. cerevisiae* as well [69–72]. Members of this group are particularly well conserved in higher metazoans, for example, being represented by the UbcH5 family in humans. They show a capacity to interact with a wide range of E3 enzymes from both HECT and RING-finger families and, depending on the specific activity of the E3 enzyme in question, are capable to facilitate mono- or polyubiquitination via K29, K48, or K63 of different proteins. This ability allows them to be involved in both proteasomal and lysosomal degradation processes [73].

#### *1. PTS1 Receptor Polyubiquitination*

The cycling PTS1 receptor Pex5p has been shown to be a substrate for Ubc4p [69–72]. Catalyzing the formation of K48-linked polyubiquitin chains, Ubc4p marks the PTS1 receptor for proteasomal degradation. In *UBC4*-deletion strains, this function can be taken over by Ubc5p [69] and Ubc1p [71], which can be explained by the fact that these enzymes share a sequence similarity of over 90% [74–76].

Polyubiquitination of Pex5p has been demonstrated to take place at the peroxisomal membrane at the end of the receptor cycle, as Pex5p has to pass the docking- and the RING-finger complexes before it can be modified. These modified receptor molecules are targeted to the 26S proteasome as polyUb-Pex5p accumulates in strains with temperature-sensitive mutants of the 19S-cap (*cim3-1*, *cim5-1*) *in vivo* [69, 72] or after treatment of lysates from wild-type cells with MG132 *in vitro* [57]. Current evidence suggests that Pex5p polyubiquitination in *S. cerevisiae* is not essential for

Pex5p function in peroxisomal protein import under physiological conditions. More likely, it seems to be a part of a quality control system that withdraws a fraction of the membrane-accumulated Pex5p that has got stuck in the export pathway by targeting it to the proteasome. Alternatively to ''quality control,'' the term RADAR (*receptor accumulation and degradation in absence of recycling*) has been suggested [27]. This mechanism is supposed to reopen binding capacities for cargo-loaded receptors at the peroxisomal membrane. However, it can be imagined that failure in the removal of Pex5p designated for disposal would affect matrix protein import. In fact, while deficiency in Ubc4p or Ubc5p alone can be compensated, the deletion of both leads to a partial import defect of peroxisomal matrix proteins [69].

Polyubiquitinated species of Pex5p accumulate in *S. cerevisiae* cells lacking ''late'' import pathway peroxins Pex1p, Pex4p, Pex6p, Pex15p, and Pex22p, making them detectable even without the inhibition of the proteasome. Apparently, the efficiency of proteasomal disposal of the receptor varies among different species drastically. While *Sc*Pex5p is heavily polyubiquitinated in mutants that are supposed to block receptor recycling, single deletions in the same set of proteins result in a dramatic decrease of the Pex5p steady state concentration in human cells [49, 77], plants [78], *Hansenula polymorpha* [79–81], or *Pichia pastoris* [53]. The phenomenon of polyUb-Pex5p accumulation in *S. cerevisiae* facilitated the discovery and functional characterization of Pex5p ubiquitination and provides the long-sought for explanation for the instability of Pex5p in human cell lines [77].

### 2. *PTS2 Coreceptor Polyubiquitination*

Ubiquitination plays an important role in regulation of the PTS2 receptor pathway. In contrast to Pex5p, which facilitates PTS1 protein import on its own, the PTS2 receptor Pex7p is associated with auxiliary proteins. *S. cerevisiae* possesses the two redundant coreceptors Pex18p and Pex21p [23].

While the PTS1 receptor Pex5p is a stable protein with a life span of over 5 h under wild-type conditions [71], the PTS2 coreceptors Pex18p and Pex21p exhibit a half-time of only about 10 min [70]. This rapid turnover may be closely connected to the physiological function of these proteins. Their degradation takes place in the 26S proteasome and is mediated by Ubc4p and Ubc5p [70]. In *ubc4*Δ*ubc5*Δ mutant cells, Pex18p is still polyubiquitinated by Ubc1p, but the efficiency of this process is reduced (Platta, unpublished data). Interestingly, Pex18p accumulates in mutants of the membrane-standing components of the peroxisomal protein import machinery, for example *pex13*Δ, *pex2*Δ or *pex1*Δ. This stabilization and

accumulation is reverted if additionally *PEX7* is disrupted, which identifies Pex7p as a stabilizing factor for Pex18p at the peroxisomal membrane. Based on coimmunoprecipitation studies, apparently not all Pex7p forms a complex with Pex18p, while the entire pool of Pex18p seems to be saturated with Pex7p [70]. These results indicate that the Ubc4p-mediated turnover of Pex18p is associated with its normal function during matrix protein import rather than being an abortive degradation.

The family of PTS2 coreceptors is evolutionary divergent. While *S. cerevisiae* contains Pex18p/Pex21p and human as well as plant cells harbor the Pex5L, most other yeasts and fungi have Pex20p-like proteins. The Pex18p/Pex21p orthologue Pex20p displays interesting parallels to the PTS1 receptor Pex5p in terms of its Ubc-based regulation. In contrast to Pex18p and Pex21p, which exhibit a turnover rate of 10 min, Pex20p of *H. polymorpha* is stable for at least 60 min [26] and is not destabilized in the absence of Pex7p in *P. pastoris* [27]. A further similarity to Pex5p is that Pex20p of *P. pastoris* is polyubiquitinated with K48-branched chains in the absence of either Pex4p, Pex1p, or Pex6p [27].

### *3. Conserved Regulatory Mechanism*

Although the degree of functional regulation of the PTS receptors in general may differ on the cell biological level, the basic biochemical recognition signals and mechanisms are the same.

It has been known that the C-terminal half of Pex5p, which contains the TPR-motifs, serves as binding region for the PTS1-cargo proteins, while the N-terminal half contains the peroxisomal targeting function, docking-complex interaction sites made of diaromatic pentapeptide repeats (Wxxx[F/Y]), and the yet not fully defined putative export signal within the first 20 amino acids [82]. The PTS2 receptor complex seems to be functionally divided in a comparable manner in that Pex7p binds the PTS2-cargo and the coreceptors seem to contain the structural requirements for the membrane-bound steps of receptor cycle [82–85]. One striking example is a chimeric protein consisting of Pex18p (without its Pex7p-binding site) fused to the TPR domains of Pex5p, which is still able to mediate PTS1 protein import [85]. Amino acid sequence alignments of the N-termini of Pex5p, Pex18p, Pex21p, and Pex20p proteins reveal that the very first lysine residue of each receptor is evolutionary highly conserved between these molecules. Experimental evidence has been provided from studies of *H. polymorpha* Pex5p [86] and *P. pastoris* Pex20p [27] that this lysine residue indeed is the acceptor of the polyubiquitin chain. Taking these findings together, an evolutionary conserved biochemical basis for Ubc4p-mediated PTS receptor ubiquitination can be defined.

## B. Function of Pex4p/Ubc10p in Receptor Recycling

Pex4p/Ubc10p is a ubiquitin-conjugating enzyme essential for peroxisomal biogenesis, which defines it as the only ubiquitin-conjugating enzyme known to be indispensable for the biogenesis of an organelle. The molecular function of this enzyme has been a mystery since its initial description in 1992 [50].

### *1. Pex4p in Peroxisomal Biogenesis*

Pex4p is anchored to the peroxisomal membrane via its interaction to the membrane-integrated Pex22p in yeasts and plants [51, 52]. This interaction is needed for Pex4p function as *pex22*Δ strains are characterized by the complete mislocalization of Pex4p to the cytosol and they share a similar phenotype like *pex4*Δ cells.

Pex4p contains the catalytically relevant active site Cys residue of ubiquitin-conjugating enzymes within the core Ubc fold [50, 87]. Site directed mutagenesis of the Cys residue results in a loss of Pex4p activity as the point mutant exhibited the same phenotype as the *pex4*Δ deletion strain, while it was still attached to peroxisomes [50]. This Cys residue is needed to form a thioester bond to ubiquitin, as this Ub-Pex4p conjugate could be detected *in vivo* under nonreducing conditions only and was disrupted under reducing conditions or in a Cys to Ala point mutated Pex4p [87]. Loss of Pex4p ubiquitin-conjugating activity is closely connected to an import defect of peroxisomal matrix proteins. This argues for a direct role of Pex4p-mediated ubiquitination in receptor cycle during matrix protein import. Indeed, Pex5p stability has been found to depend on the presence and activity of Pex4p as the Pex5p level drops significantly in PEX4 or PEX22 mutants in humans and different yeasts, with the exception of *S. cerevisiae* where polyubiquitinated Pex5p accumulates at the peroxisomal membrane [69, 72]. Instability in *pex* mutants was also observed in *P. pastoris* in the case of Pex20p [27], which was used as a tool to examine the sequence of Pex5p interactions at the peroxisomal membrane [53]. This epitasis study placed Pex1p and Pex6p downstream of the RING-finger complex and found Pex4p and Pex22p to act even later at the very end of the receptor cycle, as double deletions of Pex1p and Pex4p displayed a *pex1*Δ-like phenotype. Similar experiments in *S. cerevisiae* showed a mixed phenotype (namely a combined polyubiquitin pattern of *pex1*Δ and *pex4*Δ) and indicated the Pex5p interaction sequence to be branched spatially after the RING complex ([69] and unpublished data). Studies in *pex4*Δ strains of *H. polymorpha* found residual amounts of Pex5p to accumulate inside of peroxisomes, indicative for a functional connection of Pex4p and receptor release. The mislocalization of PTS1 matrix proteins to the cytosol in a *PEX4* deletion strain can partially be

restored by a massive overexpression of Pex5p [80]. This observation strongly suggests that it is not the physical interaction between Pex4p and Pex5p but the Pex5p-specific activity of this Ubc which is supposed to be a regulatory device involved in protein import into peroxisomes.

### 2. *Model I: Monoubiquitination*

The finding that the deletion of a ubiquitin-conjugating enzyme results in a specific polyubiquitination and degradation of its potential substrates argues for a ubiquitin-dependent, physiological, and nonproteolytic role of this modification. Interestingly, Pex5p has been reported to be monoubiquitinated at two different lysine residues in wild-type cells of *S. cerevisiae* [71]. Ubc4p is not required for Pex5p monoubiquitination, which has also been demonstrated to be the case for Ubc1p, Ubc5p, and Ubc8p, raising the attractive possibility that Pex4p may be the responsible E2 enzyme. Like the Ubc4p-dependent polyubiquitination, the monoubiquitination event takes place after docking and is located at or after the RING-finger complex. Pex5p is only transiently modified, as the monoubiquitinated Pex5p forms can only be detected by treatment with thiol-alkylating reagent *N*-ethylmaleimide (NEM), supposed to inhibit deubiquitinating enzymes, which are Cys proteases in yeasts. As monoUb-Pex5p is exclusively found at the peroxisomal membrane, it is proposed to be deubiquitinated under wild-type conditions prior or during export back to the cytosol. These results gave rise to the still hypothetical idea that Pex4p monoubiquitinates the PTS1 receptor to prime it for release to the cytosol [39].

### 3. *Model II: Polyubiquitination*

Two studies report data that argue for a polyubiquitination activity of Pex4p. Kiel and coworkers [86] noticed that obstruction of polyubiquitination by overexpression of Ub(K48R) affects PTS1 matrix protein import in the methylotrophic yeast *H. polymorpha*. Under these conditions, Pex5p is ubiquitinated and rapidly degraded by the proteasome. To test whether Ub(K48R) interferes with the function of *Hp*Pex4p, the *pex4*Δ deletion strain was complemented by massive overexpression of Pex5p as described previously [80] and in addition, Ub(K48R) was overexpressed in parallel. Interestingly, the Pex5p overexpression did result in a compensation of the Ub(K48R) effect, which was considered as evidence for a Pex4p-dependent polyubiquitination of an unknown substrate involved in receptor recycling [86]. The important role of K48-branched polyubiquitin chains for peroxisomal biogenesis has also been noticed in *P. pastoris* [27]. Overexpression of Ub(K48R) in *P. pastoris* results in a peroxisome biogenesis defect and induces accumulation, polyubiquitination, and degradation of Pex20p.

Thus, overexpression of Ub(K48R) mimics the characteristics of a defect in receptor export which might be explained by the assumption that polyubiquitination (1) might provide some kind of ''rescue signal'' under conditions when the physiological export signal for the receptors is not functional. As a result, recycling might be hampered so that the membrane has to be cleared of cargo-unloaded receptors by the proteasome in order to open binding capacities for new cargo-loaded receptors; (2) might trigger the constitutive degradation of a yet unknown protein which functions as a repressor of receptor export; or (3) might represent a physiological process concerning known parts of the export machinery itself: it has been demonstrated in other systems that E3- as well as E2 enzymes undergo autoubiquitination reactions *in vivo* [88–91].

Another study concerns Pex4p and Pex22p from *Arabidopsis thaliana* and discusses the potential role of Pex4p-mediated polyubiquitination as a regulator of matrix protein composition [51]. Glyoxysomes are specialized microbodies that function in early seedling development and are converted to peroxisomes in leafs when photosynthesis is initiated. Isocitrate lyase is a glyoxysome-specific enzyme that usually is not anymore present in leaf peroxisomes. Interestingly, this protein remains stable in peroxisomes in the absence of Pex4p and Pex22p, indicating that the proteins may be important during the remodeling of peroxisome matrix contents as glyoxysomes transition to leaf peroxisomes.

Basically, both models can be true, as Pex4p may interact with different E3 enzymes to either mono- or polyubiquitinate their substrate(s). Studies in Pex4p-affected mutants have been extremely valuable to explore the functional context of this Ubc. However, taking in account that the Pex4p/Pex22p unit is part of the peroxisomal protein import machinery, which is a multiprotein complex also consisting of docking-, RING-, and AAA-subcomplexes [47, 92, 93], it might be difficult to distinguish direct from indirect effects. For future research, it therefore will be of significant importance to establish *in vitro* assays to unequivocally prove Pex5p or other proteins to be substrates for ubiquitination reactions mediated by Pex4p. This will also require the identification of the corresponding E3 enzymes for which the peroxisomal RING-finger peroxins are the best candidates.

### C. Peroxisomal RING-Finger Proteins as Putative Ubiquitin-Ligase-Complex

The RING-finger motif was first identified in the protein encoded by the *Really Interesting New Gene 1* by Freemont *et al.* [94] and has ever since been implicated in mediating protein–protein interactions of different

kinds. In recent years, it has become evident that most if not all RING-finger-containing proteins have ubiquitin-protein ligase (isopeptidase) activity and act as E3 enzymes [95, 96]. The canonical RING-finger consensus sequence has been defined as Cys-$X_2$-Cys-$X_{9–39}$-Cys-$X_{1–3}$- His-$X_{2–3}$-Cys/His-$X_2$-Cys-$X_{4–48}$-Cys-$X_2$-Cys, where X stands for any amino acid residue. Like other Cys-rich motifs, the RING-finger binds $Zn^{2+}$ ions through its conserved Cys and His residues but is set apart from the rest by its unique ''cross-brace'' arrangement of their two Zn-coordination sites, where Cys1/Cys2 and Cys5/Cys6 bind the first and Cys3/His4 and Cys7/Cys8 bind the second. $Zn^{2+}$ ion. The majority of RING-finger proteins fall into two subclasses, RING-HC and RING-H2, depending on the presence of a Cys or a His in the fifth $Zn^{2+}$-coordination site [97].

The three peroxins Pex2p, Pex10p, and Pex12p are peroxisomal integral membrane proteins and possess cytosolically exposed RING domains in their C-termini. All belong to the RING-HC family, but only in Pex10p both $Zn^{2+}$-coordination sites are well conserved. Pex2p and Pex12p contain substitutions for the conserved Cys- and His residues in the second $Zn^{2+}$-coordination site. The zinc-binding capacity of Pex10p was investigated and the protein has been proven to coordinate zinc with its RING domain [98]. This zinc-dependent activity of the peroxisomal RING-fingers could provide the basis for the general zinc requirement of matrix protein import [99]. The RING-finger peroxins are essential for peroxisomal biogenesis in all species analyzed [100–105]. They interact with each other by forming a unique trimeric RING-finger complex and are also capable of interaction with Pex5p [48, 100, 105–107]. More than one decade after the identification of the first RING peroxin [108], the molecular function of the peroxisomal RING-finger complex is still a matter of debate. While some studies demonstrate that the RING peroxins are required for Pex5p import in an *in vitro* system [58], others find Pex5p [48, 49, 58] or the PTS2-coreceptor Pex20p [27] to accumulate inside the peroxisomal lumen in cells with disrupted RING complex. Direct evidence for E3 ligase activity of the RING peroxins is still missing. It is important to note, however, that both mono- and polyubiquitination of Pex5p depend on the presence of the RING-finger peroxins [69, 71, 72]. Additionally, a Pex22p-dependent interaction between Pex4p and Pex10p has been observed in the split ubiquitin system [109]. Thus, Pex4p and Ubc4p might be recruited to the peroxisomal membrane and exhibit their E2-activity in concert with the putative E3-ligase peroxins in order to mono- or polyubiquitinate the PTS receptors.

## V. The AAA Family ATPases

Pex1p is one of the defining members of the large AAA family of enzymes [10, 110]. AAA proteins are characterized by the presence of 200- to 250-amino acid ATP-binding domains that contain Walker A and B motifs. AAA proteins themselves belong to the superfamily of P-loop NTPases [111].

### A. Function and Structure of AAA-Type ATPases

AAA proteins have an N-terminal non-ATPase domain that is followed by either one or two AAA domains (D1 and D2). In some proteins with two AAA domains, both are evolutionarily well conserved (like in Cdc48p/97). In others, either the D2 domain (like in Pex1p and Pex6p) or the D1 domain (in Sec18p/NSF) is better conserved in evolution (Figure 21.2) The classical AAA has been expanded by inclusion of a number of more distantly related cellular regulators and termed $AAA^+$ family of ATPases [112]. $AAA^+$ proteins are involved in protein degradation, membrane fusion, DNA replication, microtubule dynamics, disassembly of protein complexes, and protein aggregates [111]. AAAs are mechanoenzymes that manipulate the structure of substrate proteins and thereby unfold them or disassemble protein complexes.

The physiologically active form of these enzymes often is a homohexamer. The hexameric enzymes have an overall shape that resembles a double-ring with a central pore that might be involved in substrate processing. In the hexameric configuration, the ATP-binding site is positioned at the interface between the subunits. On ATP binding and hydrolysis, AAA enzymes undergo conformational changes in the AAA domains as well as in the N-domains. These motions can be transmitted to substrate protein.

AAA proteins are not restricted to eukaryotes. Prokaryotes have AAA which combine chaperone with proteolytic activity, like the ClpAPS complex, which mediates protein degradation and recognition in *Escherichia coli*. The basic recognition of proteins by AAAs is thought to occur through unfolded domains in the substrate protein [113]. In HslU, a bacterial ClpX/ClpY homologue of the Hsp100 family of $AAA^+$ proteins, the N- and C-terminal subdomains move toward each other when nucleotides are bound and hydrolyzed. The terminal domains are most distant in the nucleotide-free state and closest in the ADP-bound state. Thereby the opening of the central cavity is affected [114, 115].

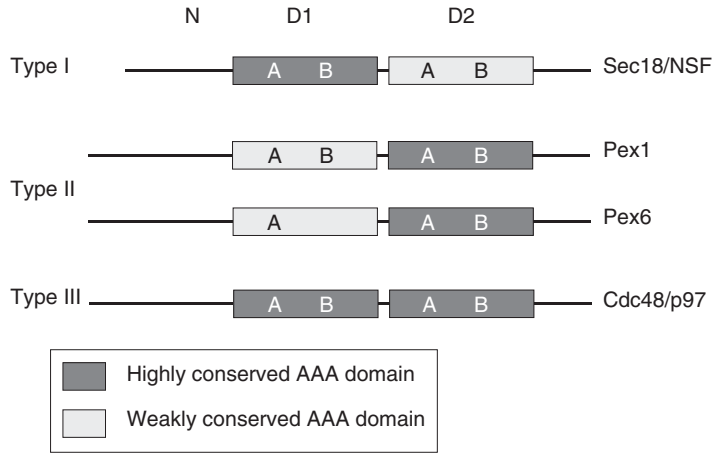



FIG. 21.2. Domain structure of AAA protein family members. Pex1p and Pex6p as well as Sec18p and Cdc48p are members of the AAA family of ATPases. These AAA proteins share a common structural organization as they exhibit an N-terminal domain (N) and two AAA ATPase domains (D1 and D2). The AAA domains contain an ATP-binding site (Walker A) and a site for ATP-hydrolysis (Walker B). Depending on the degree of conservation of the AAA domains, these AAA ATPases can be divided into three groups. Sec18p is an example for type I with the first AAA domain (D1) being better conserved than the second one (D2). Pex1p and Pex6p belong to type II as their second domain is the one that is highly conserved. It is interesting to note that Pex6p has no ATP-hydrolysis site in its first AAA cassette. Cdc48p stands for type III, where both domains are well conserved. In many cases, the N-domain of the AAA proteins proved to be capable to interact with adaptor proteins.

## B. AAA ATPases in Protein Transport

### *1. ER-Associated Protein Degradation*

The AAA-type ATPase Cdc48p/p97 is perhaps the best-studied AAA protein. Misfolded secretory proteins are exported from the ER and degraded by the ER-associated degradation pathway (ERAD) [116]. Substrate retrotranslocation and extraction are assisted by the Cdc48p (Ufd1p/Npl4p) complex on the cytosolic side of the membrane where the substrate is ubiquitinated by ER-based E2 and E3 enzymes before degradation by the 26S proteasome [116].

### *2. Targeting to Multivesicular Bodies*

Multivesicular bodies (MVB) are endosomal compartments that sort ubiquitinated membrane proteins by incorporating them into vesicles. This process involves the sequential action of three multiprotein complexes,

ESCRT I–III [62]. Vps4p is an AAA-type ATPase involved in this MVB sorting pathway. It had originally been identified as a ''class E'' vps (vacuolar protein sorting) mutant [117] and was subsequently shown to catalyze the dissociation of ESCRT complexes [118]. It is anchored via Vps46p to the endosomal membrane. Vps4p assembly is assisted by the conserved Vta1p protein, which regulates its oligomerization status and ATPase activity.

## VI. Pex1p and Pex6p: AAA Proteins Required for Peroxisomal Biogenesis

Two lines of research have converged into the recognition of AAA proteins as enzymes important for peroxisome biogenesis. One is the study of the energetic requirement of protein import. The other is the molecular characterization of AAA peroxins.

### A. ATP-Dependency of Matrix Protein Import

Peroxisomal matrix protein import was shown to be ATP-dependent in several *in vitro* systems. In permeablized CHO cells, import of PTS1 proteins depends on ATP, but not on GTP or a membrane potential [119]. As the peroxins Pex1p and Pex6p are the only ATPases among the known peroxins, they were considered as being responsible for the bulk ATP requirement of peroxisomal matrix protein import. The described ATP requirement may comprise several steps in the function of the AAA peroxins.

1. Mutations in the AAA peroxin genes represent the most frequent cause of human peroxisomal biogenesis disorders (PBDs) [120]. Cells with defective Pex1p or Pex6p mislocalize peroxisomal matrix proteins to the cytosol [10, 121, 122]. As Pex1p and Pex6p interact ATP-dependently with each other [54, 81, 123, 124], it is important to note that the most common cause of Zellweger syndrome is a disruption of the Pex1p–Pex6p interaction caused by a point mutation in the conserved ATP-binding site of Pex1p [120, 125].
2. Based on two-hybrid data, it has been discussed that ATP hydrolysis in the conserved domain of Pex6p contributes to the disassembly of the Pex6p–Pex15p complex [55], which indicates that the AAA peroxins interact dynamically with Pex15p.
3. *In vitro* experiments indicated that ATP is needed predominantly for the recycling of the PTS1 receptor Pex5p rather than for its insertion into the

membrane [41, 126]. The *in vitro* reconstitution of the complete Pex5p cycle revealed that ATP binding and hydrolysis in the conserved domains of both Pex1p and Pex6p were needed for this reaction [57]. The binding and consumption of ATP may result in conformational changes that could generate the driving force to pull the receptor out of the membrane.

4. Additionally to their involvement in matrix protein import, Pex1p and Pex6p also play a role in peroxisome membrane fusion. In the yeast *Yarrowia lipolytica*, six peroxisomal membrane subforms have been identified that in a multistep pathway assemble into mature peroxisomes [127]. The fusion of early forms in this pathway requires ATP and the AAA peroxins, which are heterogeneously distributed over these subforms [127].

### B. Structural Characterization of the AAA Peroxins

The domain structure of Pex1p and Pex6p is similar to other AAA proteins: N-D1-D2 (Figure 21.2). The second AAA domain, D2, is better conserved than the first. The mouse Pex1p N-domain is the only part of AAA peroxins for which X-ray structural information is available [128]. It consists of two structurally independent lobes separated by a shallow groove (Figure 21.3). The structure is strikingly similar to the N-domains of other AAA proteins with two ATPase domains that have been solved: Sec18p, NSF, p97, and the archaeal Cdc48p/p97 homologue VAT [129–132]. In spite of the low degree in sequence similarity, the cleft between the subdomains is structurally conserved. The structure looks similar to the cleft in NSF, which is a binding site for α-SNAP. While the adaptor protein for the N-terminal domain of Pex1p is still elusive, Pex6p has been demonstrated to interact with the peroxisomal membrane protein Pex15p in yeast [55] and Pex26p in human cells [56] via the N-terminus. Pex15p/Pex26p is the membrane recruitment factor for the cytosolic AAA complex. We have little information yet on the oligomerization status of the AAA peroxins. They are thought to build up hexameric structures [92], but it is unclear weather they form homo- or heterohexamers.

### C. Similarities of the Peroxisomal Import Machinery with ERAD Components

There is striking similarity of the peroxisomal import machinery that we portrait in this chapter with a special focus on Pex4p and the AAA peroxins on one side and the ERAD pathway on the other side

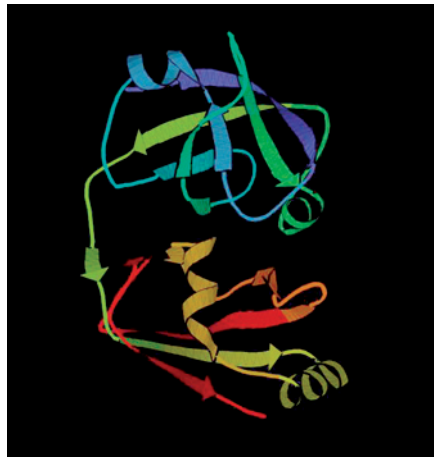


FIG. 21.3. Structure of the N-terminal domain (amino acids 3–180) of the AAA protein PEX1 from mouse. The polypeptide is folded into two globular subdomains. The N-terminal lobe folds as double-$\chi$-barrel and the C-terminal loop as a $\beta$-barrel. In spite of the low similarity (less than 10%) with other N-domains of AAA proteins, the overall structure is remarkably similar to the N-domains of Cdc48p/p97 or NSF, suggesting that these structures have a common function, for example, in substrate binding. Image was generated form Protein Data Bank (PDB) structure 1WLF using Kinemage (Duke University, United States).

(Figure 21.4). Both pathways depend on the ubiquitination cascade starting with the general ubiquitin-activating enzyme E1 (Uba1p). Both systems include two Ubcs that are recruited to their destination via a membrane anchor (Ubc7p-Cue1p/Pex4p-Pex22p). Involved are integral RING-finger proteins Doa10p and Hrd1p in ERAD, while peroxisomes rely on the trimeric RING-finger complex Pex2p-Pex10p-Pex12p. The AAA ATPase Cdc48p/p97 as well as the AAA peroxin complex Pex1p-Pex6p provide the driving force for protein extraction from the membranes. Furthermore, the proteasome is involved in both systems. The resemblance in terms of protein equipment has been backed up by two studies on the evolutionary basis of this similarity [133, 134]. Both studies concluded that peroxisomes are of eukaryotic origin with few ancestral protein motifs in prokaryotes. While abandoning the suggestion that peroxisomes might be endosymbionts, the topological question remains open. ERAD transports proteins out of the endomembrane compartment; while in peroxisome biogenesis, matrix proteins are imported. If, however, ubiquitinated ERAD substrates are equated

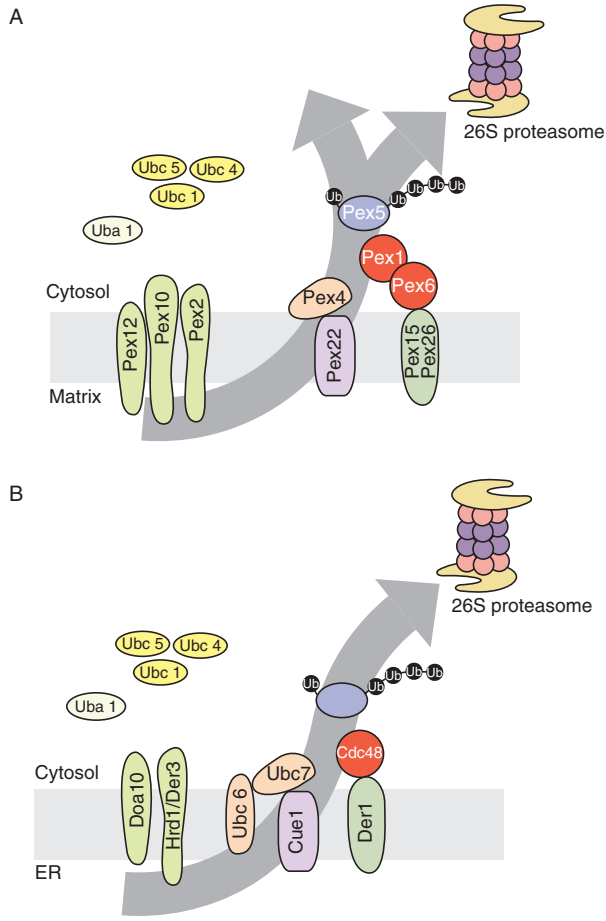



FIG. 21.4 Similarities between PTS receptor recycling during peroxisomal matrix protein import and ERAD. The schematic representation of the Pex5-recycling pathway (left) and ERAD (right) displays similarities in terms of protein composition. Both contain integral RING-finger proteins that are proven E3 enzymes (Doa10p, Hrd1p) or have putative E3 activity (Pex2p, Pex10p, Pex12p). The peroxisomal E2 enzyme Pex4p is anchored via Pex22p, whereas Ubc7p is attached to the ER membrane through Cue1p. Additionally, Ubc7p interacts with the membrane integrated Ubc6p. The members of the Ubc4p/Ubc5-pUbc1p family participate in both pathways, which also is the case for Uba1p, which charges the E2 enzymes with activated ubiquitin. Each system includes membrane-anchored AAA complexes. The AAA ATPase Cdc48p interacts with the integral ER-membrane protein Der1p, whereas the AAA peroxin Pex6p binds to Pex15p in *S. cerevisiae* or Pex26p in human cells. Additionally, Pex6p forms a complex with the second AAA peroxin, Pex1p. Both AAA peroxins are required for dislocation of the PTS1 receptor Pex5 back to the cytosol. This is reminiscent of the function of Cdc48p, which translocates polyubiquitinated proteins

with the peroxisomal import receptor, both processes can be understood as a ubiquitin-dependent protein dislocation from an endomembrane system (Figure 21.4).

## VII. Receptor Ubiquitination: A Link Between Pex4p, AAA Peroxins, and Protein Transport?

This chapter summarizes the basic experimental evidence concerning the functional roles of Pex4p and the AAA peroxins in Pex5p recycling and matrix protein import in order to combine and discuss them in a unified model (Figure 21.5).

Peroxisomal matrix protein import is an energy-dependent reaction process as hydrolysis of ATP is needed for protein translocation [119, 135]. This energy-consuming step was further characterized to involve the receptor cycle of Pex5p [49]. Studies in a permeabilized cell system of human fibroblasts provided first evidence that Pex5p accumulated reversibly under conditions when protein transport was blocked due to the absence of ATP [49]. Detailed *in vitro* studies revealed that the binding and translocation of Pex5p itself is ATP-independent, while the ATP-dependent step concerns the export of Pex5p back to the cytosol [41]. Pex14p-associated Pex5p exposes the majority of its mass into the peroxisomal lumen, suggesting that translocation of cargo proteins into the matrix of the organelle occurs concomitantly with the formation of the Pex5p–Pex14p membrane complex and is succeeded by the ATP-dependent export reaction [41, 126, 136, 137]. The ATP-dependent export of Pex5p might be the rate-limiting step in the protein import process. Supposedly, the binding capacities for Pex5p at the peroxisomal membrane are nearly saturated under wild-type conditions so that the cargo-free receptor has to be removed from the membrane in order to keep the flow of protein import going. The most simplified idea is that one cargo-free receptor has to leave in order to enable a cargo-charged receptor to enter the peroxisomal matrix. The identity of the ATPase required for Pex5p export remained a matter of debate until recently *in vitro* systems in *S. cerevisiae* [57] and human cells [58] identified the peroxisomal AAA ATPases Pex1p and Pex6p as the motor proteins of Pex5p export (Figure 21.5).

from the ER lumen to the cytosol for proteasomal disposal. This part of the pathway is branched in peroxisomes. While polyubiquitinated receptors are supposed to be degraded by the proteasome, monoubiquitination may function as signal for Pex5p recycling in order to facilitate another round of matrix protein import.

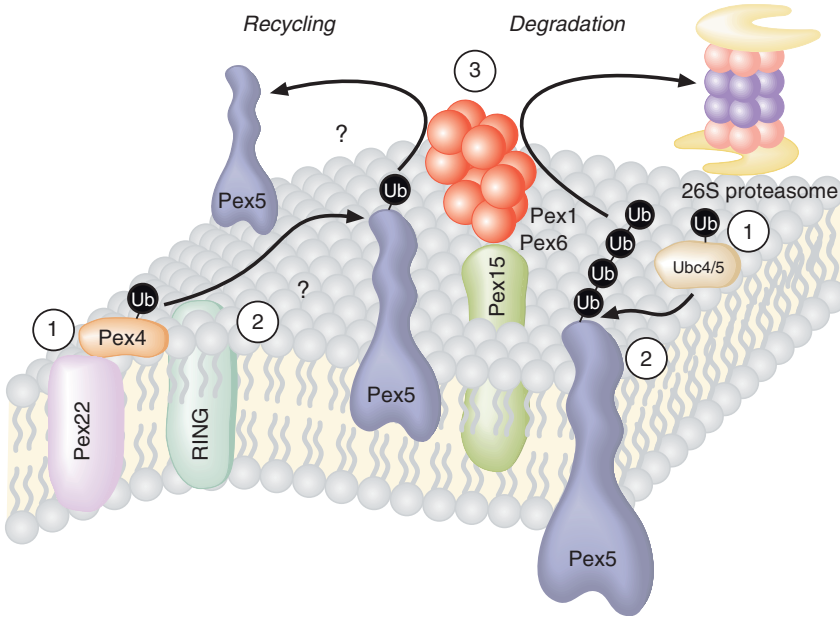



FIG. 21.5. Model for Pex5p ubiquitination and retrotranslocation. The cycling PTS1 receptor Pex5p has to be exported back to the cytosol after release of the cargo protein into the peroxisomal lumen. At this point of its cycle, Pex5p behaves like an integral membrane protein. Presumably three enzymatic activities, localized at the peroxisomal membrane, are required for Pex5p release: (1) ubiquitin-conjugation, (2) ubiquitin-protein isopeptide ligation, and (3) ATPase activity. Members of the Ubc4p/Ubc5p/Ubc1p family function in polyubiquitination of Pex5p. This process requires the RING-finger peroxins Pex2p, Pex10p, and Pex12p as well which may function as E3 enzymes. The polyubiquitinated forms of Pex5p can be recognized by the AAA peroxins Pex1p and Pex6p that are anchored to the peroxisomal membrane via Pex15p. PolyUb-Pex5p is extracted from the membrane in an AAA peroxin and ATP-dependent manner and is degraded by the 26S proteasome. Pex5p destined for recycling is supposed to be modified by monoubiquitination. This process also requires the RING-finger peroxins and is proposed to be mediated directly by the peroxisomal Ubc Pex4p. Monoubiquitinated Pex5p is supposed to be dislocated by the AAA peroxins, deubiquitinated, and thereby made available for further rounds of matrix protein import.

The ATP-binding and hydrolysis sites of the second AAA domain of Pex1p and Pex6p proved responsible for the observed energy dependence of the export process [57]. Furthermore, membrane-integrated Pex5p was disassembled from Pex14p and released to the cytosol, indicating that the integral Pex5p population is a target for the AAA-ATPases [57].

The exact mechanism in terms of substrate recognition by the AAA peroxins is still not known. The N-terminus of membrane-integrated Pex5p is exposed to the cytosol and was demonstrated to be required for *in vitro* export of human Pex5p [82]. Accumulation of an N-terminally truncated version of Pex20p at the peroxisomal membrane points to similar conditions for the PTS2 coreceptors [27]. The first 20 amino acids of the PTS receptors contain a conserved motif of unknown function (Cys-$X_n$-Asn-(Ala/Gly)-(Leu/Ala)), which could act as the putative binding site for Pex1p or Pex6p. However, to date no direct interaction of Pex5p with Pex1p or Pex6p was detected, despite their obvious formation of a complex at the peroxisomal membrane [57, 58, 92]. Thus, we also have to consider that this interaction is regulated or mediated by a third factor, leading to the more general question of how the AAA peroxins can distinguish Pex5p forms destined for dislocation from cargo-charged Pex5p species. Current evidence draws the attention to the activity of the ubiquitin-conjugating enzyme Pex4p, which is required for matrix protein import [50] and might be involved in the recycling of Pex5p [53, 69, 72, 80] and Pex20p [27]. Taking further in account that the peroxisomal RING-finger complex consisting of Pex2p, Pex10p, and Pex12p may function as putative E3-ligase complex needed for protein import and Pex5p release, the dislocation of the PTS1 receptor may be linked to ubiquitination processes. Based on the observation that Pex5p is monoubiquitinated in wild-type cells [71], the model of a Pex4p- and Pex10p-mediated monoubiquitination of the PTS1 receptor with the purpose of this modification is to prime Pex5p for efficient export mediated by the AAA peroxins has been discussed [39]. However, the experimental evidence for this modification being required for Pex5p export is still missing. It is known that ubiquitin-based signals can be recognized by proteins harboring UBDs [60]. Based on this idea, the interaction between Pex5p and Pex1p/Pex6p could be mediated by ubiquitin. This would postulate that the AAA peroxins contain UBDs or associate with adaptor proteins that can bind ubiquitin. Another possibility is based on the finding that the N-terminal half of Pex5p is a natively unfolded domain [138]. The attachment of ubiquitin to substrates can induce conformational changes within the modified protein and make formally hidden binding sites accessible [60]. This mode of interaction is also discussed for the AAA-ATPase Cdc48p/p97. While it is clear that the adaptor complexes Ufd1-Npl4 or Ubx proteins bind polyubiquitin chains, the N-terminus of the AAA ATPase itself is able to recognize chains as well as nonmodified segments of its substrates [139, 140]. Thus, monoubiquitination of Pex5p may alter its folding state and enable a direct interaction with the AAA peroxins not via but depending on ubiquitin. As a consequence of receptor

export, ubiquitin has to be cleaved off after or during dislocation by one of the many deubiquitinating enzymes [68].

The importance of an optimal PTS receptor release is also supposed to be the reason for the above-mentioned existence of a quality control system at the peroxisomal membrane [27, 69, 72, 138], which leads to the receptor polyubiquitination and proteasomal degradation in mutants lacking peroxins involved in PTS receptor export [69, 71, 72]. A physiological function for polyubiquitination in peroxisomal biogenesis can be postulated from the ''transient pore model'' [39]. In this context, it is thinkable that the pore forming Pex5p population represents a dead-end structure that is withdrawn from the recycling process and has to be removed by the ubiquitin-proteasome system. Alternatively, polyubiquitination may substitute for the putative monoubiquitin-based export signal for receptor release under yet to define conditions. A similar effect has been demonstrated in the case of p53 export from the nucleus [65]. Although mono- and polyubiquitination target p53 to different fates, at least one of them has to be present for dislocation to the cytosol.

Summarizing the current evidence, one can draw the conclusion that the energy dependence of peroxisomal protein import is caused by the cycle of the PTS receptors. Retrotranslocation of the receptors is the energy- and thus most likely the rate-limiting step of matrix protein import. This energy dependence can be separated into two groups of enzymatic activities: on the one hand, receptor ubiquitination by Pex4p, as ubiquitin has to be activated by E1 before Pex4p can be charged; and on the other hand, ATP hydrolysis in the conserved AAA domain of Pex1p and Pex6p in order to pull the primed Pex5p out of the membrane.

A model is emerging in which the previously disparate roles of Pex4p and the AAA peroxins are combined in a concerted reaction sequence. For future research, it will be a challenge to elucidate how ATP-dependent receptor dislocation is mechanistically linked to import of folded proteins.

## References

1. van den Bosch, H., Schutgens, R.B., Wanders, R.J., and Tager, J.M. (1992). Biochemistry of peroxisomes. *Annu Rev Biochem* 61:157–197.
2. Eckert, J.H., and Erdmann, R. (2003). Peroxisome biogenesis. *Rev Physiol Biochem Pharmacol* 147:75–121.
3. Wanders, R.J. (2004). Metabolic and molecular basis of peroxisomal disorders: a review. *Am J Med Genet* 126A:355–375.
4. Rhodin, J. (1954). Correlation of ultrastructural organization and function in normal and experimentally changed peroxisomal convoluted tubule cells of the mouse kidney. *Dissertation, karolinska institutet Aktiebolaget Godvil, Stockholm.*

5. DeDuve, C., and Baudhuin, P. (1966). Peroxisomes (microbodies and related particles). *Physiol Rev* 46:323–357.
6. Parsons, M., Furuya, T., Pal, S., and Kessler, P. (2001). Biogenesis and function of peroxisomes and glycosomes. *Mol Biochem Parasitol* 115:19–28.
7. Hiltunen, J.K., Mursula, A.M., Rottensteiner, H., Wierenga, R.K., Kastaniotis, A.J., and Gurvitz, A. (2003). The biochemistry of peroxisomal beta-oxidation in the yeast *Saccharomyces cervisiae*. *FEMS Microbiol Rev* 27:35–64.
8. Baker, A., and Sparkes, I.A. (2005). Peroxisome protein import: some answers, more questions. *Curr Opin Plant Biol* 8:640–647.
9. Erdmann, R., Veenhuis, M., Mertens, D., and Kunau, W.-H. (1989). Isolation of peroxisome-deficient mutants of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 86:5419–5423.
10. Erdmann, R., Wiebel, F.F., Flessau, A., Rytka, J., Beyer, A., Fröhlich, K.U., and Kunau, W.-H. (1991). *PAS1*, a yeast gene required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. *Cell* 64:499–510.
11. Lazarow, P.B. (1993). Genetic approaches to study peroxisome biogenesis. *Trends Cell Biol* 157:105–132.
12. Distel, B., Erdmann, R., Gould, S.J., Blobel, G., Crane, D.I., Cregg, J.M., Dodt, G., Fujiki, Y., Goodman, J.M., Just, W.W., Kiel, A.K.W., Kunau, W.-H., *et al.* (1996). A unified nomenclature for peroxisome biogenesis factors. *J Cell Biol* 135:1–3.
13. van der Zand, A., Braakman, I., Geuze, H.J., and Tabak, H.F. (2006). The return of the peroxisome. *J Cell Sci* 119:989–994.
14. Lazarow, P.B., and Fujiki, Y. (1985). Biogenesis of peroxisomes. *Annu Rev Cell Biol* 1:489–530.
15. Brul, S., Westerveld, A., Strijland, A., Wanders, R.J., Schram, A.W., Heymans, H.S., Schutgens, R.B., van den Bosch, H., and Tager, J.M. (1988). Genetic heterogeneity in the cerebrohepatorenal (Zellweger) syndrome and other inherited disorders with a generalized impairment of peroxisomal functions. A study using complementation analysis. *J Clin Invest* 81:1710–1715.
16. Stewart, M.Q., Esposito, R.D., Gowani, J., and Goodman, J.M. (2001). Alcohol oxidase and dihydroxyacetone synthase, the abundant peroxisomal proteins of methylotrophic yeasts, assemble in different cellular compartments. *J Cell Sci* 114:2863–2868.
17. Titorenko, V.I., Nicaud, J.M., Wang, H., Chan, H., and Rachubinski, R.A. (2002). Acyl-CoA oxidase is imported as a heteropentameric, cofactor-containing complex into peroxisomes of *Yarrowia lipolytica*. *J Cell Biol* 156:481–494.
18. Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J., and Subramani, S. (1989). A conserved tripeptide sorts proteins to peroxisomes. *J Cell Biol* 108:1657–1664.
19. McCollum, D., Monosov, E., and Subramani, S. (1993). The *pas8* mutant of *Pichia pastoris* exhibits the peroxisomal protein import deficiencies of Zellweger syndrome cells: the Pas8 protein binds to the COOH-terminal tripeptide peroxisomal targeting signal, and is a member of the TPR protein family. *J Cell Biol* 121:761–774.
20. Swinkels, B.W., Gould, S.J., Bodnar, A.G., Rachubinski, R.A., and Subramani, S. (1991). A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *EMBO J* 10:3255–3262.
21. Marzioch, M., Erdmann, R., Veenhuis, M., and Kunau, W.-H. (1994). *PAS7* encodes a novel yeast member of the WD-40 protein family essential for import of 3-oxoacyl-CoA thiolase, a PTS2-containing protein, into peroxisomes. *EMBO J* 13:4908–4918.
22. Zhang, J.W., and Lazarow, P.B. (1995). *PEB1* (*PAS7)* in *Saccharomyces cerevisiae* encodes a hydrophilic, intra-peroxisomal protein that is a member of the WD repeat family and is essential for the import of thiolase into peroxisomes. *J Cell Biol* 129:65–80.

23. Purdue, P.E., Yang, X., and Lazarow, P.B. (1998). Pex18p and Pex21p, a novel pair of related peroxins essential for peroxisomal targeting by the PTS2 pathway. *J Cell Biol* 143:1859–1869.
24. Titorenko, V.I., Smith, J.J., Szilard, R.K., and Rachubinski, R.A. (1998). Pex20p of the yeast *Yarrowia lipolytica* is required for the oligomerization of thiolase in the cytosol and for its targeting to the peroxisome. *J Cell Biol* 142:403–420.
25. Sichting, M., Schell-Steven, A., Prokisch, H., Erdmann, R., and Rottensteiner, H. (2003). Pex7p and Pex20p of *Neurospora crassa* function together in PTS2-dependent protein import into peroxisomes. *Mol Biol Cell* 14:810–821.
26. Otzen, M., Wang, D., Lunenborg, M.G., and van der Klei, I.J. (2005). *Hansenula polymorpha* Pex20p is an oligomer that binds the peroxisomal targeting signal 2 (PTS2). *J Cell Sci* 118:3409–3418.
27. Leon, S., Zhang, L., McDonald, W.H., Yates, J., IIIrd., Cregg, J.M., and Subramani, S. (2006). Dynamics of the peroxisomal import cycle of PpPex20p: ubiquitin-dependent localization and regulation. *J Cell Biol* 172:67–78.
28. Braverman, N., Dodt, G., Gould, S.J., and Valle, D. (1998). An isoform of Pex5p, the human PTS1 receptor, is required for the import of PTS2 proteins into peroxisomes. *Hum Mol Genet* 7:1195–1205.
29. Otera, H., Harano, T., Honsho, M., Ghaedi, K., Mukai, S., Tanaka, A., Kawai, A., Shimizu, N., and Fujiki, Y. (2000). The mammalian peroxin Pex5pL, the longer isoform of the mobile peroxisome targeting Signal (PTS) type 1 transporter, translocates the Pex7p-PTS2 protein complex into peroxisomes via its initial docking site, Pex14p. *J Biol Chem* 275:21703–21714.
30. Woodward, A.W., and Bartel, B. (2005). The arabidopsis peroxisomal targeting signal type 2 receptor PEX7 is necessary for peroxisome function and dependent on PEX5. *Mol Biol Cell* 16:573–583.
31. Lee, J.R., Jang, H.H., Park, J.H., Jung, J.H., Lee, S.S., Park, S.K., Chi, Y.H., Moon, J.C., Lee, Y.M., Kim, S.Y., Yun, D.J., Cho, M.J., *et al.* (2006). Cloning of two splice variants of the rice PTS1 receptor, OsPex5pL and OsPex5pS, and their functional characterization using pex5-deficient yeast and Arabidopsis. *Plant J* 47:457–466.
32. Legakis, J.E., and Terlecky, S.R. (2001). PTS2 protein import into mammalian peroxisomes. *Traffic* 2:252–260.
33. Huhse, B., Rehling, P., Albertini, M., Blank, L., Meller, K., and Kunau, W.-H. (1998). Pex17p of *Saccharomyces cerevisiae* is a novel peroxin and component of the peroxisomal protein translocation machinery. *J Cell Biol* 140:49–60.
34. Girzalsky, W., Rehling, P., Stein, K., Kipper, J., Blank, L., Kunau, W.-H., and Erdmann, R. (1999). Involvement of Pex13p in Pex14p localization and peroxisomal targeting signal 2 dependent protein import into peroxisomes. *J Cell Biol* 144:1151–1162.
35. Saidowsky, J., Dodt, G., Kirchberg, K., Wegner, A., Nastainczyk, W., Kunau, W.H., and Schliebs, W. (2001). The Di-aromatic pentapeptide repeats of the human peroxisome import receptor PEX5 are separate high-affinity binding sites for the peroxisomal membrane protein PEX14. *J Biol Chem* 276:34524–34529.
36. Niederhoff, K., Meindl-Beinker, N.M., Kerssen, D., Perband, U., Schäfer, A., Schliebs, W., and Kunau, W.H. (2005). Yeast Pex14p possesses two functionally distinct Pex5p and one Pex7p binding sites. *J Biol Chem* 280:35571–35578.
37. Williams, C., van den Berg, M., and Distel, B. (2005). *Saccharomyces cerevisiae* Pex14p contains two independent Pex5p binding sites, which are both essential for PTS1 protein import. *FEBS Lett* 579:3416–3420.

38. Urquhart, A.J., Kennedy, D., Gould, S.J., and Crane, D.I. (2000). Interaction of Pex5p, the type 1 peroxisome targeting signal receptor, with the peroxisomal membrane proteins Pex14p and Pex13p. *J Biol Chem* 275:4127–4136.
39. Erdmann, R., and Schliebs, W. (2005). Peroxisomal matrix protein import: the transient pore model. *Nat Rev Mol Cell Biol* 6:738–742.
40. Baker, A., and Sparkes, I.A. (2005). Peroxisome protein import: some answers, more questions. *Curr Opin Plant Biol* 8:640–670.
41. Gouveia, A.M., Reguenga, C., Oliveira, M.E., Sa-Miranda, C., and Azevedo, J.E. (2000). Characterization of peroxisomal Pex5p from rat liver: Pex5p in the Pex5p-Pex14p membrane complex is a transmembrane protein. *J Biol Chem* 275:32444–32451.
42. Dammai, V., and Subramani, S. (2001). The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* 105:187–196.
43. Kunau, W. (2001). Peroxisomes: the extended shuttle to the peroxisome matrix. *Curr Biol* 11:R659–R662.
44. Nair, D.M., Purdue, P.E., and Lazarow, P.B. (2004). Pex7p translocates in and out of peroxisomes in *Saccharomyces cerevisiae*. *J Cell Biol* 167:599–604.
45. Wang, D., Visser, N.V., Veenhuis, M., and Van Der Klei, I.J. (2003). Physical interactions of the peroxisomal targeting signal 1-receptor, Pex5p, studied by fluorescence correlation spectroscopy. *J Biol Chem* 278:43340–43345.
46. Zhang, L., Leon, S., and Subramani, S. (2006). Two independent pathways traffic the intraperoxisomal peroxin PpPex8p into peroxisomes: mechanism and evolutionary implications. *Mol Biol Cell* 17:690–699.
47. Agne, B., Meindl, N.M., Niederhoff, K., Einwächter, H., Rehling, P., Sickmann, A., Meyer, H.E., Girzalsky, W., and Kunau, W.H. (2003). Pex8p. An intraperoxisomal organizer of the peroxisomal import machinery. *Mol Cell* 11:635–646.
48. Chang, C.C., Warren, D.S., Sacksteder, K.A., and Gould, S.J. (1999). PEX12 interacts with PEX5 and PEX10 and acts downstream of receptor docking in peroxisomal matrix protein import. *J Cell Biol* 147:761–774.
49. Dodt, G., and Gould, S.J. (1996). Multiple *PEX* genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor. *J Cell Biol* 135:1763–1774.
50. Wiebel, F.F., and Kunau, W.-H. (1992). The PAS2 protein essential for peroxisome biogenesis is related to ubiquitin-conjugating enzymes. *Nature* 359:73–76.
51. Zolman, B.K., Monroe-Augustus, M., Silva, I.D., and Bartel, B. (2005). Identification and functional characterization of arabidopsis PEROXIN4 and the interacting protein PEROXIN22. *Plant Cell* 17:3422–3435.
52. Koller, A., Snyder, W.B., Faber, K.N., Wenzel, T.J., Rangell, L., Keller, G.A., and Subramani, S. (1999). Pex22p of *Pichia pastoris*, essential for peroxisomal matrix protein import, anchors the ubiquitin-conjugating enzyme, Pex4p, on the peroxisomal membrane. *J Cell Biol* 146:99–112.
53. Collins, C.S., Kalish, J.E., Morrell, J.C., McCaffery, J.M., and Gould, S.J. (2000). The peroxisome biogenesis factors Pex4p, Pex22p, Pex1p, and Pex6p act in the terminal steps of peroxisomal matrix protein import. *Mol Cell Biol* 20:7516–7526.
54. Tamura, S., Shimozawa, N., Suzuki, Y., Tsukamoto, T., Osumi, T., and Fujiki, Y. (1998). A cytoplasmic AAA family peroxin, Pex1p, interacts with Pex6p. *Biochem Biophys Res Commun* 245:883–886.

55. Birschmann, I., Stroobants, A.K., Van Den Berg, M., Schäfer, A., Rosenkranz, K., Kunau, W.H., and Tabak, H.F. (2003). Pex15p of *Saccharomyces cerevisiae* provides a molecular basis for recruitment of the AAA Peroxin Pex6p to peroxisomal membranes. *Mol Biol Cell* 14:2226–2236.
56. Matsumoto, N., Tamura, S., and Fujiki, Y. (2003). The pathogenic peroxin Pex26p recruits the Pex1p-Pex6p AAA ATPase complexes to peroxisomes. *Nat Cell Biol* 5:454–460.
57. Platta, H.W., Grunau, S., Rosenkranz, K., Girzalsky, W., and Erdmann, R. (2005). Functional role of the AAA peroxins in dislocation of the cycling PTS1 receptor back to the cytosol. *Nat Cell Biol* 7:817–822.
58. Miyata, N., and Fujiki, Y. (2005). Shuttling mechanism of peroxisome targeting signal type 1 receptor Pex5: ATP-independent import and ATP-dependent export. *Mol Cell Biol* 25:10822–10832.
59. Pickart, C.M. (2001). Mechanisms underlying ubiquitination. *Annu Rev Biochem* 70:503–533.
60. Hicke, L., Schubert, H.L., and Hill, C.P. (2005). Ubiquitin-binding domains. *Nat Rev Mol Cell Biol* 6:610–621.
61. Haglund, K., and Dikic, I. (2005). Ubiquitylation and cell signaling. *EMBO J* 24:3353–3359.
62. Hurley, J.H., and Emr, S.D. (2006). The ESCRT complexes: structure and mechanism of a membrane-trafficking network. *Annu Rev Biophys Biomol Struct* 35:277–298.
63. Sigismund, S., Polo, S., and Di Fiore, P.P. (2004). Signalling through monoubiquitination. *Curr Top Microbiol Immunol* 286:149–185.
64. Miyase, S., Tateishi, S., Watanabe, K., Tomita, K., Suzuki, K., Inoue, H., and Yamaizumi, M. (2005). Differential regulation of Rad18 through Rad6-dependent mono- and polyubiquitination. *J Biol Chem* 280:515–524.
65. Li, M., Brooks, C.L., Wu-Baer, F., Chen, D., Baer, R., and Gu, W. (2003). Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. *Science* 302:1972–1975.
66. Robzyk, K., Recht, J., and Osley, M.A. (2000). Rad6-dependent ubiquitination of histone H2B in yeast. *Science* 287:501–504.
67. Friedberg, E. (2006). Reversible monoubiquitination of PCNA: a novel slant on regulating translesion DNA synthesis. *Mol Cell* 22:150–152.
68. Amerik, A.Y., and Hochstrasser, M. (2004). Mechanism and function of deubiquitinating enzymes. *Biochim Biophys Acta* 1695:189–207.
69. Platta, H.W., Girzalsky, W., and Erdmann, R. (2004). Ubiquitination of the peroxisomal import receptor Pex5p. *Biochem J* 384:37–45.
70. Purdue, P.E., and Lazarow, P.B. (2001). Pex18p is constitutively degraded during peroxisome biogenesis. *J Biol Chem* 276:47684–47689.
71. Kragt, A., Voorn-Brouwer, T.M., Van den Berg, M., and Distel, B. (2005). The *Saccharomyces cerevisiae* peroxisomal import receptor Pex5p is monoubiquitinated in wild type cells. *J Biol Chem* 280:7867–7874.
72. Kiel, J.A., Emmrich, K., Meyer, H.E., and Kunau, W.H. (2005). Ubiquitination of the PTS1 receptor, Pex5p, suggests the presence of a quality control mechanism during peroxisomal matrix protein import. *J Biol Chem* 280:1921–1930.
73. Weissman, A.M. (2001). Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol* 2:169–178.
74. Kuan, J., and Saier, M.H. (1993). The mitochondrial carrier family of transport proteins: structural, functional, and evolutionary relationships. *Crit Rev Biochem Mol Biol* 28:209–233.
75. Seufert, W., McGrath, J.P., and Jentsch, S. (1990). UBC1 encodes a novel member of an essential subfamily of yeast ubiquitin-conjugating enzymes involved in protein degradation. *EMBO J* 9:4535–3541.

76. Seufert, W., and Jentsch, S. (1990). Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. *EMBO J* 9:543–550.
77. Yahraus, T., Braverman, N., Dodt, G., Kalish, J.E., Morrell, J.C., Moser, H.W., Valle, D., and Gould, S.J. (1996). The peroxisome biogenesis disorder group 4 gene, PXAAA1, encodes a cytoplasmic ATPase required for stability of the PTS1 receptor. *EMBO J* 15:2914–2923.
78. Zolman, B.K., and Bartel, B. (2004). An Arabidopsis indole-3-butyric acid-response mutant defective in PEROXIN6, an apparent ATPase implicated in peroxisomal function. *Proc Natl Acad Sci USA* 101:1786–1791.
79. Rabouille, C., Levine, T.P., Peters, J.M., and Warren, G. (1995). An NSF-like ATPase, p97, and NSF mediate cisternal regrowth from mitotic Golgi fragments. *Cell* 82:905–914.
80. van der Klei, I.J., Hilbrands, R.E., Kiel, J.A.K.W., Rasmussen, S.W., Cregg, J.M., and Veenhuis, M. (1998). The ubiquitin-conjugating enzyme Pex4p of *Hansenula polymorpha* is required for efficient functioning of the PTS1 import machinery. *EMBO J* 17:3608–3618.
81. Kiel, J.A., Hilbrands, R.E., van der Klei, I.J., Rasmussen, S.W., Salomons, F.A., van der Heide, M., Faber, K.N., Cregg, J.M., and Veenhuis, M. (1999). *Hansenula polymorpha* Pex1p and Pex6p are peroxisome-associated AAA proteins that functionally and physically interact. *Yeast* 15:1059–1078.
82. Costa-Rodrigues, J., Carvalho, A.F., Gouveia, A.M., Fransen, M., Sa-Miranda, C., and Azevedo, J.E. (2004). The N-terminus of the peroxisomal cycling receptor, Pex5p, is required for redirecting the peroxisome-associated peroxin back to the cytosol. *J Biol Chem* 279:46573–46579.
83. Dodt, G., Warren, D., Becker, E., Rehling, P., and Gould, S.J. (2001). Domain mapping of human PEX5 reveals functional and structural similarities to *S. cerevisiae* Pex18p and Pex21p. *J Cell Biol* 276:41769–41781.
84. Einwächter, H., Sowinski, S., Kunau, W.H., and Schliebs, W. (2001). *Yarrowia lipolytica* Pex20p, *Saccharomyces cerevisiae* Pex18p/Pex21p and mammalian Pex5pL fulfil a common function in the early steps of the peroxisomal PTS2 import pathway. *EMBO Rep* 2:1035–1039.
85. Schäfer, A., Kerssen, D., Veenhuis, M., Kunau, W.H., and Schliebs, W. (2004). Functional similarity between the peroxisomal PTS2 receptor binding protein Pex18p and the N-Terminal Half of the PTS1 receptor Pex5p. *Mol Cell Biol* 24:8895–8906.
86. Kiel, J.A., Otzen, M., Veenhuis, M., and van der Klei, I.J. (2005). Obstruction of polyubiquitination affects PTS1 peroxisomal matrix protein import. *Biochim Biophys Acta* 1745:176–186.
87. Crane, D.I., Kalish, J.E., and Gould, S.J. (1994). The *Pichia pastoris PAS4* gene encodes a ubiquitin-conjugation enzyme required for peroxisome assembly. *J Biol Chem* 269:21835–21844.
88. Walter, J., Urban, J., Volkwein, C., and Sommer, T. (2001). Sec61p-independent degradation of the tail-anchored ER membrane protein Ubc6p. *EMBO J* 20:3124–3131.
89. Banerjee, A., Gregori, L., Xu, Y., and Chau, V. (1993). The bacterially expressed yeast CDC34 gene product can undergo autoubiquitination to form a multiubiquitin chain-linked protein. *J Biol Chem* 268:5668–5675.
90. Nuber, U., Schwarz, S.E., and Scheffner, M. (1998). The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate. *Eur J Biochem* 254:643–649.
91. Hetfeld, B.K., Helfrich, A., Kapelari, B., Scheel, H., Hofmann, K., Guterman, A., Glickmann, M., Schade, R., Kloetzel, P.M., and Dubiel, W. (2005). The zinc finger of the CSN-associated deubiquitinating enzyme USP15 is essential to rescue the E3 ligase Rbx1. *Curr Biol* 15:1217–1221.

92. Rosenkranz, K., Birschmann, I., Grunau, S., Girzalsky, W., Kunau, W.-H., and Erdmann, R. (2006). Functional association of the AAA-complex and the peroxisomal importomer. *FEBS J* 273:3804–3815.
93. Hazra, P.P., Suriapranata, I., Snyder, W.B., and Subramani, S. (2002). Peroxisome remnants in pex3Delta cells and the requirement of Pex3p for interactions between the peroxisomal docking and translocation subcomplexes. *Traffic* 3:560–574.
94. Freemont, P.S., Hanson, I.M., and Trowsdale, J. (1991). A novel cysteine-rich sequence motif. *Cell* 64:483–484.
95. Joazeiro, C.A., and Weissman, A.M. (2000). RING finger proteins: mediators of ubiquitin ligase activity. *Cell* 102:549–552.
96. Ardley, H.C., and Robinson, P.A. (2005). E3 ubiquitin ligases. *Essays Biochem* 41:15–30.
97. Freemont, P.S. (2000). RING for destruction? *Curr Biol* 10:R84–R87.
98. Kalish, J.E., Theda, C., Morrell, J.C., Berg, J.M., and Gould, S.J. (1995). Formation of the peroxisome lumen is abolished by loss of *Pichia pastoris* Pas7p, a zinc-binding integral membrane protein of the peroxisome. *Mol Cell Biol* 15:6406–6419.
99. Terlecky, S.R., Legakis, J.E., Hueni, S.E., and Subramani, S. (2001). Quantitative analysis of peroxisomal protein import *in vitro*. *Exp Cell Res* 263:98–106.
100. Albertini, M., Girzalsky, W., Veenhuis, M., and Kunau, W.-H. (2001). Pex12p of *Saccharomyces cerevisiae* is a component of a multi-protein complex essential for peroxisomal matrix protein import. *Eur J Cell Biol* 80:257–270.
101. Warren, D.S., Morrell, J.C., Moser, H.W., Valle, D., and Gould, S.J. (1998). Identification of *PEX10*, the gene defective in complementation group 7 of the peroxisome-biogenesis disorders. *Am J Hum Genet* 63:347–359.
102. Fan, J., Quan, S., Orth, T., Awai, C., Chory, J., and Hu, J. (2005). The arabidopsis PEX12 gene is required for peroxisome biogenesis and is essential for development. *Plant Physiol* 139:231–239.
103. Tan, X., Waterham, H.R., Veenhuis, M., and Cregg, J.M. (1995). The *Hansenula polymorpha PER8* gene encodes a novel peroxisomal integral membrane protein involved in proliferation. *J Cell Biol* 128:307–319.
104. Schumann, U., Gietl, C., and Schmid, M. (1999). Sequence analysis of a cDNA encoding Pex10p, a zinc-binding peroxisomal integral membrane protein from *Arabidopsis thaliana*. *Plant Physiol* 119:1147–1147.
105. Krazy, H., and Michels, P.A. (2006). Identification and characterization of three peroxins-PEX6, PEX10 and PEX12-involved in glycosome biogenesis in *Trypanosoma brucei*. *Biochim Biophys Acta* 1763:6–17.
106. Reguenga, C., Oliveira, M.E., Gouveia, A.M., Sa-Miranda, C., and Azevedo, J.E. (2001). Characterization of the mammalian peroxisomal import machinery: Pex2p, Pex5p, Pex12p and Pex14p are subunits of the same protein assembly. *J Biol Chem* 273:29935–29942.
107. Okumoto, K., Abe, I., and Fujiki, Y. (2000). Molecular anatomy of the peroxin Pex12p: RING finger domain is essential for the Pex12p function and interacts with the peroxisome targeting signal type 1-receptor Pex5p and a RING peroxin, Pex10p. *J Biol Chem* 275:25700–25710.
108. Tsukamoto, T., Miura, S., and Fujiki, Y. (1991). Restoration by a 35K membrane protein of peroxisome assembly in a peroxisome-deficient mammalian cell mutant. *Nature* 350:77–81.

109. Eckert, J.H., and Johnsson, N. (2003). Pex10p links the ubiquitin conjugating enzyme Pex4p to the protein import machinery of the peroxisome. *J Cell Sci* 116:3623–3634.
110. Kunau, W.-H., Beyer, A., Franken, T., Götte, K., Marzioch, M., Saidowsky, J., Skaletz-Rorowski, A., and Wiebel, F.F. (1993). Two complementary approaches to study peroxisome biogenesis in *Saccharomyces cerevisiae*: forward and reversed genetics. *Biochimie* 75:209–224.
111. Hanson, P.I., and Whiteheart, S.W. (2005). AAA + proteins: have engine, will work. *Nat Rev Mol Cell Biol* 6:519–529.
112. Neuwald, A.F., Aravind, L., Spouge, J.L., and Koonin, E.V. (1999). $AAA^{+}$: a class of chaperone-like ATPases associated with the assembly, operation and disassembly of protein complexes. *Genome Res* 9:27–43.
113. Prakash, S., Tian, L., Ratliff, K.S., Lehotzky, R.E., and Matouschek, A. (2004). An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 11:830–837.
114. Wang, J., Song, J.J., Franklin, M.C., Kamtekar, S., Im, Y.J., Rho, S.H., Seong, I.S., Lee, C.S., Chung, C.H., and Eom, S.H. (2001). Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure* 9:177–184.
115. Sousa, M.C., TRame, C.B., Tsuruta, H., Wilbanks, S.M., Reddy, V.S., and McKay, D.B. (2000). Crystal and solution structures of an HsIUV protease-chaperone complex. *Cell* 103:633–643.
116. Romisch, K. (2005). Endoplasmatic reticulum-associated degradation. *Annu Rev Cell Dev Biol* 21:435–456.
117. Babst, M., Sato, T.K., Banta, L.M., and Emr, S.D. (1997). Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. *EMBO J* 16:1820–1831.
118. Babst, M., Katzmann, D.J., Estepa-Sabal, E.J., Meerloo, T., and Emr, S.D. (2002). ESCRT-III: an endosome-associated heterooligomeric protein complex required for MVB sorting. *Dev Cell* 3:271–281.
119. Wendland, M., and Subramani, S. (1993). Cytosol-dependent peroxisomal protein import in a permeabilized cell system. *J Cell Biol* 120:675–685.
120. Reuber, B.E., Germain-Lee, E., Collins, C.S., Morrell, J.C., Ameritunga, R., Moser, H.W., Valle, D., and Gould, S.J. (1997). Mutations in *PEX1* are the most common cause of peroxisome biogenesis disorders. *Nat Genet* 17:445–448.
121. Spong, A.P., and Subramani, S. (1993). Cloning and characterization of *PAS5*: a gene required for peroxisome biogenesis in the methylotrophic yeast *Pichia pastoris*. *J Cell Biol* 123:535–548.
122. Nuttley, W.M., Brade, A.M., Eitzen, G.A., Veenhuis, M., Aitchison, J.D., Szilard, R.K., Glover, J.R., and Rachubinski, R.A. (1994). *PAY4*, agene required for peroxisome assembly in the yeast *Yarrowia lipolytica*, encodes a novel member of a family of putative ATPases. *J Biol Chem* 269:556–566.
123. Birschmann, I., Rosenkranz, K., Erdmann, R., and Kunau, W.-H. (2005). Structural and functional analysis of the interaction of the AAA-peroxins Pex1p and Pex6p. *FEBS J* 272:47–58.
124. Faber, K.N., Heyman, J.A., and Subramani, S. (1998). Two AAA family peroxins, PpPex1p and PpPex6p, interact with each other in an ATP-dependent manner and are associated with different subcellular membranous structures distinct from peroxisomes. *Mol Cell Biol* 18:936–943.

125. Geisbrecht, B.V., Collins, C.S., Reuber, B.E., and Gould, S.J. (1998). Disruption of a PEX1-PEX6 interaction is the most common cause of the neurologic disorders Zellweger syndrome, neonatal adrenoleukodystrophy, and infantile Refsum disease. *Proc Natl Acad Sci USA* 95:8630–8635.
126. Oliveira, M.E., Gouveia, A.M., Pinto, R.A., Sa-Miranda, C., and Azevedo, J.E. (2003). The energetics of Pex5p-mediated peroxisomal protein import. *J Biol Chem* 278:39483–39488.
127. Titorenko, V.I., and Rachubinski, R.A. (2000). Peroxisomal membrane fusion requires two AAA family ATPases, Pex1p and Pex6p. *J Cell Biol* 150:881–886.
128. Shiozawa, K., Maita, N., Tomii, K., Seto, A., Goda, N., Akiyama, Y., Shimizu, T., Shirakawa, M., and Hiroaki, H. (2004). Structure of the N-terminal domain of PEX1 AAA-ATPase: characterization of a putative adaptor-binding domain. *J Biol Chem* 279:50060–50068.
129. Yu, R.C., Jahn, R., and Brunger, A.T. (1999). NSF N-terminal domain crystal structure: models of NSF function. *Mol Cell* 4:97–107.
130. May, A.P., Misura, K.M., Whiteheart, S.W., and Weis, W.I. (1999). Crystal structure of the amino-terminal domain of N-ethylmaleimide-sensitive fusion protein. *Nat Cell Biol* 1:175–182.
131. Coles, M., Diercks, T., Liermann, J., Groger, A., Rockel, B., Baumeister, W., Koretke, K.K., Lupas, A., Peters, J., and Kessler., H. (1999). The solution structure of VAT-N reveals a ''missing-link'' in the evolution of complex enzymes form a simple beta-alpha-beta-beta element *Curr Biol* 9:1158–1168.
132. Zhang, X., Shaw, A., Bates, P.A., Newman, R.H., Gowen, B., Orlova, E., Gorman, M.A., Kondo, H., Dokurno, P., Lally, J., Leonard, G., Meyer, H. *et al.* (2000). Structure of the AAA ATPase p97. *Mol Cell* 6:1473–1484.
133. Gabaldon, T., Snel, B., van Zimmeren, F., Hemrika, W., Tabak, H.F., and Huynen, M.A. (2006). Origin and evolution of the peroxisomal proteome. *Biol Direct* 1:1–14.
134. Schluter, A., Fourcade, S., Ripp, R., Mandel, J.L., Poch, O., and Pujol, A. (2006). The evolutionary origin of peroxisomes: an ER-peroxisome connection. *Mol Biol Evol* 23:838–845.
135. Imanaka, T., Small, G.M., and Lazarow, P.B. (1987). Translocation of acyl-CoA oxidase into peroxisomes requires ATP hydrolysis but not a membrane potential. *J Cell Biol* 105:2915–2922.
136. Gouveia, A.M., Guimaraes, C.P., Oliveira, M.E., Reguenga, C., Sa-Miranda, C., and Azevedo, J.E. (2002). Characterization of the peroxisomal cycling receptor, Pex5p, using a cell-free *in vitro* import system. *J Biol Chem* 278:226–232.
137. Gouveia, A.M., Guimaraes, C.P., Oliveira, M.E., Reguenga, C., Sa-Miranda, C., and Azevedo, J.E. (2003). Characterization of the peroxisomal cycling receptor Pex5p import pathway. *Adv Exp Med Biol* 544:213–220.
138. Carvalho, A.F., Costa-Rodrigues, J., Correia, I., Costa Pessoa, J., Faria, T.Q., Martins, C.L., Fransen, M., Sa-Miranda, C., and Azevedo, J.E. (2006). The N-terminal half of the peroxisomal cycling receptor Pex5p is a natively unfolded domain. *J Mol Biol* 356:864–875.
139. Ye, Y., Meyer, H.H., and Rapoport, T.A. (2003). Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. *J Cell Biol* 162:71–84.
140. Thoms, S. (2002). Cdc48 can distinguish between native and non-native proteins in the absence of cofactors. *FEBS Lett* 520:107–110.

# *Author Index*

Numbers in regular font are reference numbers and indicate that an author's work is referred to although the name is not cited in the text. Numbers in italics refer to the page numbers on which the complete reference appears.

**B**

**C**

**D**

**G**

**H**

**I**



**J**

**K**

**L**

**M**

**O**

**P**

**Q**



**R**

**S**

**T**

**U**



**V**

**X**



**Y**

**Z**

This page intentionally left blank

# *Index*

**M**

**Q**

**T**