Chapter 1
Biogenesis of membrane bound respiratory complexes in *Escherichia coli*

Claire E. Price and Arnold J.M. Driessen

*Escherichia coli* is one of the preferred bacteria for studies on the energetics and regulation of respiration. Respiratory chains consist of primary dehydrogenases and of terminal reductases or oxidases linked by quinones. There is a multitude of primary dehydrogenases, terminal reductases and quinones and so depending on the growth conditions, a large variability in the composition of respiratory chains is observed. In order to assemble this complex arrangement of protein complexes, synthesis of the subunits occurs in the cytoplasm followed by assembly in the cytoplasm and/or membrane, the incorporation of metal or organic cofactors and the anchoring of the complex to the membrane. In the case of exported metalloproteins, synthesis, assembly and incorporation of metal cofactors must be completed before translocation across the cytoplasmic membrane. Coordination data on these processes is, however, scarce. In this review, we discuss the various processes respiratory proteins must undergo for correct assembly and functional coupling to the electron transport chain in *E. coli*. Targeting to and translocation across the membrane together with cofactor synthesis and insertion are discussed in a general manner followed by a review of the coordinated biogenesis of individual respiratory enzyme complexes. Lastly, we address the supramolecular organization of respiratory enzymes into supercomplexes and their localization to specialized domains in the membrane.

Accepted for publication in BBA - Molecular Cell Research
INTRODUCTION

*Escherichia coli* is one of the preferred bacteria for studies on the energetics and regulation of respiration. Respiratory chains consist of primary dehydrogenases and of terminal reductases or oxidases linked by quinones. There is a multitude of primary dehydrogenases, terminal reductases and quinones and so depending on the growth conditions, a large variability in the composition of respiratory chains is observed.

To date 15 primary dehydrogenases and 10 terminal reductases have been identified in *E. coli* (488). Terminal reductases and dehydrogenases are expressed under specific conditions depending on the availability of oxygen and other electron acceptors. Therefore the number of enzymes that are actually able to interact is restricted (487). There are also energetic restraints precluding the cooperation of some dehydrogenases and terminal reductases (81,475).

Bacterial respiratory chains are composed of a variety of electron transport constituents, for example flavoproteins, iron-sulfur (Fe-S) proteins, quinones and cytochromes. The differential transport of electrons and protons through the cytoplasmic membrane leads to the formation of an electrochemical proton gradient across the membrane which can be used to drive ATP formation via the F$_1$F$_0$ ATP synthase or other energy requiring processes such as active transport of substrates in and out of the cell.

When oxygen is present it is the preferred terminal electron acceptor while alternative electron acceptors such as nitrate, fumarate, dimethylsulfoxide (DMSO) and trimethylamine N-oxide (TMAO) can be used in the absence of oxygen. The electron chain is branched into quinone-reducing and quinone-oxidizing branches as illustrated in Figure 1. In the reducing branch, electrons are transferred from substrates such as succinate, lactate, formate, H$_2$ and NADH to quinones. The quinone-oxidizing branch contains proteins that oxidize quinol such as the cytochrome c reductase and the terminal oxidases. Terminal oxidases are oxidoreductases that are capable of oxygen reduction during aerobic respiration or the reduction of electron acceptors other than oxygen during anaerobic respiration. A number of respiratory complexes consist of cytoplasmically active enzymes that are anchored to the cytoplasmic membrane by one or two subunits which are often $b$-type cytochromes. The membrane subunits connect the cytoplasmic redox reactions with electron transport to or from a quinone/quinol pool. For example fumarate reductase (217), formate dehydrogenase N (231) and nitrate reductase A.
Introduction

(43) all share this structure. These enzymes can form subcomplexes of cytoplasmic subunits even in the absence of the membrane anchoring subunits and usually retain their oxidoreductase activity although this activity is uncoupled from electron transfer to or from the quinone/quinol pool. This suggests that the attachment of the enzymes to the membrane by their cytochrome b subunits is the last step in complex assembly. Other respiratory complexes consist of periplasmically-located catalytic subunits which are anchored to the cytoplasmic membrane by an integral membrane protein. The catalytic subunits must therefore be transported across the cytoplasmic membrane, often after cofactor insertion and protein folding. For example, the periplasmic nitrate reductase (224) and DMSO reductase (299) have periplasmically-located catalytic subunits. There are also isoenzymes for some electron donors such as H₂, formate, NADH and glycerol-3-phosphate and for electron acceptors such as O₂ and NO₂. Isoenzymes are completely different from one another in genetic and biochemical properties.

FIGURE 1. Simplified scheme of the aerobic and anaerobic respiration electron pathways in *E. coli*. The quinone-reducing (left) and quinol-oxidizing (right) branches with terminal oxidases reducing either O₂ or alternative electron acceptors are shown. The quinol pool consists of ubiquinone, menaquinone and demethylmenaquinone. Grey boxes indicate electron input units and black arrows the influx of reducing equivalents. The white boxes indicate terminal oxidases/reductases and the electron flux to these complexes is indicated by white arrows.
Chapter 1

Most structural genes for respiratory enzymes are organized in one operon. An exception to this rule is the formate hydrogen-lyase. Some operons, for example cyoABCD and narGHJI, contain additional genes required for maturation or processing of the enzyme or for the biosynthesis of cofactors. The products of the additional genes are not found in the mature protein (48,56,128,192,399). In Table 1 the structural components of the respiratory enzymes covered in this chapter are summarized together with the reactions catalyzed. In order to assemble this intricate arrangement of protein complexes, synthesis of the subunits occurs in the cytoplasm followed by assembly in the cytoplasm and/or membrane, the incorporation of metal or organic cofactors and the anchoring of the complex to the membrane. In the case of exported metalloproteins, synthesis, assembly and incorporation of metal cofactors must be completed before translocation across the cytoplasmic membrane. In this chapter, the various processes respiratory proteins must undergo for correct membrane insertion, assembly and functional coupling to the electron transport chain in E. coli are discussed. Targeting to and translocation across the membrane together with cofactor synthesis and insertion are discussed in a general manner followed by a review of the coordinated biogenesis of individual respiratory enzyme complexes.

2. MEMBRANE PROTEIN BIOGENESIS IN E. coli

2.1 GENERAL SECRETORY PATHWAY

In E. coli, the majority of protein translocation across or insertion into the cytoplasmic membrane occurs via the general secretory pathway, otherwise known as the Sec system. In this pathway, ribosome-bound nascent chains are targeted by the bacterial signal recognition particle (SRP) to the translocase via the SRP receptor FtsY, in the case of integral membrane proteins, while secretory proteins are bound post-translationally by the chaperone SecB which targets the protein to the translocase while keeping it in a translocation-competent state (see (124,282) for recent reviews). Proteins secreted via the Sec translocase are synthesized with an N-terminal extension known as the signal sequence (178). The signal sequence is around 20 amino acids in length and can be subdivided into three domains: The N-domain which is the N-terminal region containing positive charges; the H-domain which is largely hydrophobic; and the C-domain at the C-terminal region which contains the signal peptide cleavage site.

The protein conducting channel of the Sec translocase consists of the
dimeric SecYEG complex (75,496). Associated with this channel is the motor protein SecA. SecA is an ATPase which drives the translocation of proteins and large polar intermembrane loops across the membrane. Through successive cycles of ATP hydrolysis, SecA drives the translocation of proteins through the protein conducting channel in a step wise manner (426,485). SecA supports translocation at a constant rate and this rate is independent of the length of the secretory protein (472).

During or shortly after translocation across the inner membrane, the signal peptide is removed by membrane-located signal peptidases. Pre-protein processing by signal peptidases is essential for bacterial growth and viability (110). The signal peptide of most translocated proteins and some inner membrane proteins is cleaved by leader peptidase, also known as signal peptidase I or LepB (106,553). The signal peptide of lipoproteins is cleaved by signal peptidase II or LspA (213,471) while a third signal peptidase, signal peptidase IV, cleaves type IV prepilins and prepilin-like components of the type II secretion machinery (360). Following preprotein cleavage, the signal peptide is degraded. First protease IV cuts the signal peptide after which the cytoplasmically-located oligopeptidase A degrades the cut portions of the signal peptide (340,341,458).

The structure of the SecYEβ translocon from Methanococcus jannaschii has been elucidated (496) as well as that of the Sec translocase complex including SecA from Thermotoga maritima (552). In the SecYEβ structure, the translocase contains one copy of each of the three subunits (Figure 2A). SecY is organized into N- and C-terminal domains each comprising of 5 transmembrane segments (TMSs). The domains are connected by the periplasmic loop between TMSs 5 and 6. The translocase can be likened to a clamshell that can open toward the lipid bilayer with the hinge located between TMSs 5 and 6. The two smaller subunits, SecE and Secβ (analagous to SecG in E. coli) are located at the periphery of the translocon and SecE may act as a brace preventing the domains of SecY from separating completely. The structure shows a cytoplasmic funnel which tapers to a close in the middle of the membrane. This constriction, or pore ring, is formed by 6 hydrophobic amino acid residues and may fit around the translocating polypeptide chain, providing a seal that restricts the movement of ions and other small molecules during protein translocation. The constriction is blocked at the periplasmic face of the pore by the plug region formed by TMS 2a. Opening of the channel would therefore require movement of the plug. Plug movement is observed upon SecA binding in the T. maritima structure (552) but the plug
TABLE 1. Membrane-associated respiratory chain proteins covered in this chapter.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Catalytic reaction</th>
<th>Expression conditions</th>
<th>Structural components</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>NADH dehydrogenase I</td>
<td>$\text{NADH} + \text{H}^+ + \text{UQ} + 4\text{H}^+_\text{out} \rightarrow \text{NAD}^+ + \text{UQH}_2 + 4\text{H}^+$</td>
<td>aerobic/anaerobic</td>
<td>NuoABC/DEF GHIJKLMN</td>
<td>(27,419) and <em>Thermus thermophilus</em> (420)</td>
</tr>
<tr>
<td>NADH dehydrogenase II</td>
<td>$\text{NADH} + \text{H}^+ \rightarrow \text{NAD}^+ + 2\text{H}^+ + 2\text{e}^-$</td>
<td>aerobic</td>
<td>Ndh</td>
<td>n.a.</td>
</tr>
<tr>
<td>Succinate dehydrogenase</td>
<td>Succinate $\rightarrow$ $2\text{H}^+ + 2\text{e}^- + \text{fumarate}$</td>
<td>aerobic</td>
<td>SdhABCD</td>
<td>(539)</td>
</tr>
<tr>
<td>Fumarate reductase</td>
<td>$2\text{H}^+ + 2\text{e}^- + \text{fumarate} \rightarrow \text{succinate}$</td>
<td>anaerobic</td>
<td>FrdABCD</td>
<td>(217)</td>
</tr>
<tr>
<td>Formate dehydrogenase N</td>
<td>$\text{HCOOH} \rightarrow \text{CO}_2 + 2\text{H}^+ + 2\text{e}^-$</td>
<td>anaerobic</td>
<td>FdnGHI</td>
<td>(231)</td>
</tr>
<tr>
<td>Formate dehydrogenase O</td>
<td>$\text{HCOOH} + \text{NAD}^+ \rightarrow \text{CO}_2 + \text{NADH}$</td>
<td>aerobic</td>
<td>FdoGHI</td>
<td>n.a.</td>
</tr>
<tr>
<td>Hydrogenase 1</td>
<td>$\text{H}_2 \rightarrow 2\text{H}^+ + 2\text{e}^-$</td>
<td>anaerobic</td>
<td>HyaABC</td>
<td>n.a.</td>
</tr>
<tr>
<td>Hydrogenase 2</td>
<td>$\text{H}_2 \rightarrow 2\text{H}^+ + 2\text{e}^-$</td>
<td>anaerobic</td>
<td>HybOABC</td>
<td>n.a.</td>
</tr>
<tr>
<td>Hydrogenase 3</td>
<td>$2\text{H}^+ + 2\text{e}^- \rightarrow \text{H}_2$</td>
<td>anaerobic</td>
<td>HycBCDEFG</td>
<td>n.a.</td>
</tr>
<tr>
<td>TMAO reductase</td>
<td>$(\text{CH}_3)_3\text{N}=\text{O} + 2\text{H}^+ + 2\text{e}^- \rightarrow (\text{CH}_3)_3\text{N} + \text{H}_2\text{O}$</td>
<td>anaerobic</td>
<td>TorAC</td>
<td><em>Shewanella massilia</em> (103,103)</td>
</tr>
<tr>
<td>DMSO reductase</td>
<td>$(\text{CH}_3)_2\text{SO} + 2\text{H}^+ + 2\text{e}^- \rightarrow (\text{CH}_3)_2\text{S} + \text{H}_2\text{O}$</td>
<td>anaerobic</td>
<td>DmsABC</td>
<td><em>Rhodobacter capsulatus</em> (299)</td>
</tr>
<tr>
<td>Nitrate reductase A</td>
<td>$\text{NO}_3^- + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{NO}_2^- + \text{H}_2\text{O}$</td>
<td>anaerobic</td>
<td>NarGHI</td>
<td>(43)</td>
</tr>
<tr>
<td>Periplasmic nitrate reductase</td>
<td>$\text{NO}_3^- + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{NO}_2^- + \text{H}_2\text{O}$</td>
<td>anaerobic</td>
<td>NapABC</td>
<td>(224)</td>
</tr>
<tr>
<td>Nitrite reductase</td>
<td>$\text{NO}_2^- + 8\text{H}^+ + 6\text{e}^- \rightarrow \text{NH}_4^+ + 2\text{H}_2\text{O}$</td>
<td>anaerobic</td>
<td>NrfABCD</td>
<td>(25)</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Reaction</td>
<td>Environment</td>
<td>Name</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Cytochrome o oxidase</td>
<td>$\text{O}_2 + 2\text{H}^+ + 2e^- \rightarrow 2\text{H}_2\text{O}$</td>
<td>aerobic</td>
<td>CyoABCD</td>
<td>(3)</td>
</tr>
<tr>
<td>Cytochrome bd oxidase</td>
<td>$\text{O}_2 + 2\text{H}^+ + 2e^- \rightarrow 2\text{H}_2\text{O}$</td>
<td>microaerobic</td>
<td>CydAB</td>
<td>n.a.</td>
</tr>
<tr>
<td>F$_{\text{Fo}}$ATP synthase</td>
<td>$3\text{ADP} + 3\text{P}<em>i + 10\text{H}^+</em>{\text{out}} \rightarrow 3\text{ATP} + 10\text{H}^+_{\text{in}}$</td>
<td>aerobic/anaerobic</td>
<td>AtpBEFAGDC</td>
<td>(71)</td>
</tr>
</tbody>
</table>

*From *E. coli* unless otherwise stated

* not available
domain still closes the pore in this conformation. The structure of the translocase complex showed one molecule of SecA bound to one copy of the SecYEG translocation channel (Figure 2B). The oligomerization state of actively translocating SecA is however still contested (114,141,293,351). Binding of SecA to the channel also generates changes in the lateral gate of SecY. The lateral gate is formed by TMS 2b, 3, 7 and 8 and is postulated to open toward the lipid bilayer for the release of TMSs. The SecA-induced movement of the TMSs leads to the formation of a “window” in the lateral gate proposed to accommodate the signal sequence once it is released from SecA. If the lateral gate is fixed by cross-linking with short spacers, SecA-mediated translocation is impaired indicating that SecA activation and lateral gate opening are coordinated events (125). In addition to plug movement, widening of the pore must occur to allow polypeptide chain translocation.

Associated with the Sec translocase is the heterotrimeric complex consisting of SecD, SecF and YajC. The exact role that the SecDFYajC complex plays in protein translocation is not clear. YajC, which forms a stable complex with SecDF, is not needed for cell viability while cells lacking SecDF are cold sensitive and defective in protein translocation (364). Protein translocation can, however, be achieved in vitro by functionally reconstituted SecYE and SecA, indicating that SecDF is not essential for translocation (76). Suggested functions for SecDF include clearance of the channel from processed signal sequences (297), regulation of the SecA catalytic cycle (133) and formation or regulation of an oligomeric SecYEG channel complex.

In addition to protein translocation across the cytoplasmic membrane, the Sec translocase functions in the insertion of the vast majority of integral inner membrane proteins. The membrane insertion of these proteins proceeds by a co-translational “threading mechanism” (111). The binding of ribosome-bound nascent chains may also regulate channel opening. It has been shown that ribosomes destabilize the closed state of the channel resulting in an increased ion conductance when a non-translating ribosome is bound to the channel (442). Ribosome binding would not prevent the separation of the two halves of the channel as ribosomes bind exclusively to cytosolic loops located in the C-terminal half of Sec61 which is an eukaryotic homologue of SecY (378). TMSs are postulated to exit and partition into the lipid bilayer via the lateral gate and the accessory protein YidC (discussed below) may play an important role in the clearance of transmembrane segments from the SecYEG channel (408,408).
2.2 MEMBRANE INSERTASE, YidC

YidC can function together with the SecYEG channel and also independently as a membrane insertase. A small subset of integral membrane proteins are targeted directly to YidC where they are integrated into the membrane in a Sec-independent manner. YidC belongs to the evolutionarily conserved Oxa1/Alb3/YidC family which evolved before the divergence of the 3 major domains of life (281,366,542). Oxa1 (Oxidase assembly) from yeast was the first member of this family to be described. It was originally identified as an essential factor for the biogenesis of respiratory complexes in the mitochondrion (31,64) more specifically the insertion of subunits of the cytochrome c oxidase and ATP synthase (15). Alb3 is located in the thylakoid membranes of Arabidopsis chloroplasts (319), and involved in the biogenesis of light harvesting complexes. Members of the family have all been implicated in membrane protein biogenesis of respiratory and
energy transducing proteins. In *E. coli* it has been shown that YidC is essential for the insertion of subunits a and c of the F_{1}:F_{0} ATP synthase (F\textsubscript{a} and F\textsubscript{c}) (499,499,544), subunit a of cytochrome *o* oxidase (CyoA) (126,494,494) and, more recently, subunit K of the NADH dehydrogenase complex I (377).

*E. coli* YidC has 5 TMSs with similarity to those of Oxa1 and an additional anchor TMS near the N-terminus. The protein has a large periplasmic domain between the first two TMSs and a N- and C-in topology. This structure is shared by all Gram negative bacteria but differs from that of Gram positive ones (226,397,489). Gram positive bacteria lack the N-terminal anchor TMS and large extracellular domain (470) (Figure 3). The large periplasmic domain of *E. coli* YidC has been crystallized and the structure has been solved to 2.5 Å (348). The domain forms a twisted β-sandwich with an α-helical linker which orientates the sandwich near the core TMS region (380). A portion of the periplasmic domain mediates an interaction with SecF (336,338,535) but is not essential for inner membrane biogenesis or cell viability (226). YidC is remarkably resilient to single amino acid substitutions, and even when TMSs 3 and 4 (numbering as in Figure 3) are swapped for unrelated TMSs, YidC activity is retained (226). The C-terminal 5 TMSs but not their specific sequences are essential for YidC activity. The TMSs of YidC are postulated to form a membrane chaperone, assisting the integration TMSs of membrane proteins (226).

Proteins are integrated into the membrane in a co-translational process but YidC lacks the C-terminal ribosome binding domain found in Oxa1p. Although a medium resolution cryo-EM structure of an *E. coli* ribosome nascent chain complex with YidC has been recently published (256), binding of YidC to ribosomes is not observed *in vitro* (164). It is not clear whether the ribosome would need to be targeted directly to YidC for protein insertion or simply to the membrane. The manner in which substrates for the Sec-independent insertion pathway are targeted to YidC and/or the membrane is at present unknown. It was generally accepted that the bacterial SRP specifically targeted proteins to the Sec translocase (112,113,491) but some reports in the literature (145,492,543) suggest that the same pathway can be used to target membrane proteins to YidC. Currently, it is unclear how the SRP pathway differentiates between membrane proteins destined to be inserted via SecYEG or YidC. Possibly, targeting to and insertion into the membrane are two separate processes and there may be, in the case of YidC-only substrates, a degree of flexibility in the targeting requirements (144,162,495).
2.3 TWIN-ARGININE TRANSPORT (Tat) PATHWAY

Respiratory proteins in which the catalytic subunits are periplasmically located, need to be translocated across the cytoplasmic membrane once folding and cofactor insertion has occurred in the cytoplasm (411). The translocation of such folded protein complexes is achieved via the Tat system. In *E. coli*, this system comprises of the TatA, B, C and E proteins (538) and is energized by the transmembrane electrochemical proton gradient, not requiring any nucleotides or other soluble factors (538). The Tat A, B and E proteins are sequence-related and TatA and TatE have overlapping functions (415) while TatB is essential for protein export (416). These 3 proteins share a similar structure with a single N-terminal TMS followed by an cytoplasmically-located amphipathic α-helix. TatC is a larger protein with 6 TMSs. Cross-linking data (63) and affinity purification experiments (115) suggest that multiple copies of each Tat component are present within each active Tat translocon.

The Tat system derives its name from the specialized N-terminal signal sequence present on its substrates. The signal sequence has a similar tripartate structure to those secreted via the Sec translocase but contains the distinctive SRRxFLK or “twin-arginine” sequence motif (38). This conserved motif has allowed the identification of 27 Tat-targeted proteins in *E. coli*, the majority being cofactor-containing enzymes (38,41,481,531). In most cases these enzymes must acquire...
prosthetic group(s) in the cytoplasm before export (411) indicating that the N-terminal “twin-arginine” motif is not the only structural determinant conferring Tat-dependence for translocation. The presence of a “Sec-avoidance” motif has been also been suggested (58).

The current model for Tat-dependent translocation is illustrated in Figure 4. TatA forms the translocation channel and under resting conditions exists separately from TatB and C. TatA complexes observed by electron microscopy were of variable sizes suggesting that different numbers of TatA monomers may combine to generate differently-sized channels for the translocation of diverse folded proteins across the membrane (182).

Nascent proteins to be translocated by the Tat system are first contacted by trigger factor, a chaperone with prolyl isomerase activity, although this protein is not involved in substrate targeting (230). Tat transport is a post-translational process and there are specific and non-specific chaperones involved. The chaperones DnaK and SlyD bind a broad range of Tat signal sequences (55,187,548) while some respiratory complexes have their own dedicated chaperones, examples of which are explored in detail in later sections of this chapter. The apoenzyme then binds its respective cofactors in the cytoplasm which in turn induces protein folding rendering the precursors Sec-incompatible (for example (390,411)). Protein transport via the Tat translocon follows. Lastly the Tat signal peptide is cleaved by signal peptidase I (538). Following translocation the TatABC complex dissociates, possibly due to the disengagement of the signal peptide.

FIGURE 4. Model for Tat-mediated protein translocation. Tat-dependent substrates acquire their cofactors in the cytoplasm and are presented to the Tat translocase folded either with or without the aid of chaperones. Binding of the Tat signal peptide to the TatBC subcomplex recruits TatA to form the active translocase. Translocation of the protein most likely occurs via a pore formed by an unknown number of TatA subunits, followed by proton translocation after which the signal peptide is cleaved by LepB and the protein is released into the periplasm. Next, the Tat translocase disassembles.
3. COFACTOR INSERTION IN RESPIRATORY PROTEINS

The incorporation of cofactors such as heme, molybdenum and nickel occur either after folding of the apoprotein into a conformation suitable for cofactor binding, or to the unfolded polypeptide and thus promoting a certain way of folding.

3.1 HEME COFACTORS

Heme is one of the cofactors most commonly associated with respiratory enzymes. Heme $b$, which is synthesized via the C5 pathway, is the precursor for the formation of hemes c, d, o and a. The reader is referred to reviews on the subject for an in-depth view on heme synthesis (32,105,343,517). In *E. coli*, only c-type cytochromes have heme bound covalently. Enzymes involved in the ligation of heme cofactors to cytochromes must be able to carry out a number of reactions. There needs to be specific recognition and selection of the apoprotein and subsequent release of the protein once heme ligation is complete. The enzymes must also bind and subsequently release the heme cofactor. They need to position the apoprotein and heme cofactor sufficiently close to each other with the correct geometry for the addition reaction to proceed. They should then catalyze the formation of the thioester bonds. Also, since membrane bound cytochromes can contain heme-cofactors on either side of the cytoplasmic membrane, heme must be translocated across the membrane. It has been postulated that all the above reactions are catalyzed by the protein complex cytochrome c heme lyase or CCHL. CCHL is a complex of proteins encoded in *E. coli* by the ccm operon. The operon encodes for 8 integral membrane proteins (CcmA-H) and is the best-characterized active heme transport system (184,469). CcmE acts as a periplasmic heme chaperone, taking heme from CcmC (148,430,431). CcmC-bound holo-CcmE can function as a heme reservoir (148) but because heme is bound with high affinity to CcmC (147), an ATP-dependent release complex consisting of CcmABC is required to release heme (as holo-CcmE). It has been proposed that holo-CcmE is the substrate for the cytochrome c synthetase (CcmF/H) for subsequent heme ligation to apocytochrome c (8,148). The function of CcmD has been a question of considerable interest (9,183,384) since it shows little conservation of the primary sequence (386). CcmD is not required for the transfer of heme to CcmE but is required, as part of the CcmABC complex, for the release of holo-CcmE from the release complex (386). There is likely another active heme transport system in *E. coli* encoded by *cydCD*. There is a high degree of sequence similarity between
Chapter 1

CydCD and ABC transporters shown to be involved in heme transport and cydCD mutants are unable to synthesize holocytochrome bd or c-type cytochromes \(369,370\).

All cytochromes other than c-type cytochromes have noncovalently bound heme which is usually bound through the axial ligation of the heme Fe with one or two nucleophilic amino acid residues in the polypeptide. Most often these residues are histidines where the nonprotonated nitrogen of the imidazole ring provides the electrons for heme Fe binding. Mutation of the histidines results in fully assembled but inactive cytochromes lacking individual heme cofactors \(476\).

No proteins have been shown to catalyze noncovalent heme insertion. CydCD proteins may translocate heme across the membrane to allow assembly of the bd-type oxidase \(369\). The cytochrome b\textsubscript{558} of the bd complex is thought to be located on the periplasmic side of the membrane \(146\) thus necessitating active transport of the heme cofactor across the cytoplasmic membrane. Heme binding occurs after membrane insertion of the subunits of the o and bd type oxidases and subunits I and II of the oxidases are integrated stably into the membrane in the absence of heme association \(273,333,382,439\).

3.2 Fe-S CLUSTERS

Proteins containing [Fe-S] clusters play essential roles in electron transport during respiration. \textit{E. coli} has ISC (Iron Sulfur Cluster) and SUF (Sulfur assimilation) systems involved in the assembly of iron and sulfur atoms into various clusters from which they can be transferred to protein acceptors (for extensive reviews see \(152,227\)). There is no specific cluster machinery for each [Fe-S] protein to mature. The ISC and SUF systems are contained in single operons and have similar biochemical activity. The \textit{isc} operon mediates the general or housekeeping pathway for [Fe-S] cluster biosynthesis while the \textit{suf} operon plays a role under stress conditions \(324,325,353,551\). There is also a third homologous operon present, the \textit{csd} operon, but it is not known under which physiological conditions the products of this operon are needed \(310\).

In the ISC system, the Fe-S cluster is assembled on IscU followed by transfer to the apoprotein \(6,7,66,165,342,490,534\). The efficiency of transfer is greatly enhanced if the molecular chaperone composed of HscAB is present \(89,134\). Deletion of IscS, which catalyzes cysteine desulfuration to alanine and sulfide, results in an overall decrease in activity of all Fe-S proteins studied and a general growth defect \(432\). The mutant strains showed a 50-fold reduction in NADH dehydrogenase I activity.
and a 5-fold reduction in the activity of succinate dehydrogenase. Also, amply overexpressed [4Fe-4S] or [2Fe-2S] ferredoxins is not observed without the concurrent overexpression of the isc and hsc gene clusters (460). In vitro it has been shown that other proteins can function to donate the sulfur atom for [Fe-S] cluster assembly but whether this applies to E. coli in vivo is unclear (150,151).

Iron has a low solubility and is toxic to the cell. It is therefore likely that iron donor proteins exist that would bind and hence solubilize the iron before delivering it to the desulfurase or scaffold proteins (365). The proposed iron donor function could be carried out by frataxins as they can bind iron and exist in oligomeric active forms (6,14,72,100,177,269,344) and have been shown to interact with scaffold proteins and desulfurases (6,177). The cytoplasmically-located frataxin, CyaY, appears to play an important but not essential role in [Fe-S] cluster biogenesis (93,278,365,545). CyaY has been shown to bind to IscS and deliver iron to the scaffold protein IscU (269).

The binding sites of proteins contain sequences of conserved arrangements of cysteines and histidines which serve as Fe complexing ligands (175). The binding sites of some Fe-S proteins are located on the periplasmic side of the membrane but it is unclear how iron and sulfur atoms would reach the periplasm to become associated with such sites.

3.3 OTHER COFACTORs

The incorporation of Cu into subunits of heme-Cu oxidases such as the o type oxidase is likely to be a late step in the assembly of the protein complexes (reviewed in (210)). E. coli contains a protein disulphide isomerase, DipZ, which is essential for the formation of mature cytochromes which may represent a link between cytochrome maturation and copper metabolism (102).

In the hyp operon, several proteins encoded by the operon are involved in the formation of the active hydrogenase rather than forming structural components of the hydrogenase itself (56). HypB for example, is a nickel and GTP-binding protein which delivers Ni to the hydrogenase apoprotein (288). Also, there is a dedicated nickel transporter encoded by nik (137,533).

Molybdenum is bound in the molybdopterin cofactor whose synthesis requires at least 5 different operons: moaABC, mobAB, moeB, mogA and modABCD which encodes for a molybdenum-specific ABC transporter (298,379,434). The binding site for the molybdopterin cofactor of nitrate reductase A is located on the cytoplasmic side of the membrane where as the periplasmic nitrate, TMAO and
DMSO reductases are all periplasmic molybdoproteins. The cofactors may therefore be incorporated in the cytoplasm before the subunits are translocated across the membrane.

In succinate dehydrogenase and fumarate reductase, flavin adenine dinucleotide (FAD) is covalently linked to the soluble polypeptides before it associates with the other subunits of the enzyme complexes (521,529). The mechanism by which FAD is covalently linked to the subunits is not known.

3.4 FOLDING
Folding of polypeptides is tightly connected with cofactor binding and oligomerization of the subunits. It is not clear whether protein folding influences cofactor incorporation and oligomerization in multisubunit complexes or vice versa. The general chaperones GroEL and DnaK may play a role in cytochrome biogenesis and CcmE and H may be periplasmic cytochrome c-specific chaperones. In contrast to molecular chaperones which prevent incorrect protein-protein interactions, folding catalysts catalyze defined steps in protein folding. Protein disulphide isomerase-folding catalysts play an important role in cytochrome biogenesis as they catalyze the formation or cleavage of disulphide bonds. They are found in the cytoplasm and periplasm. For example, cytoplasmic thioredoxin and thioredoxin reductase (29,156) prevent disulphide bond formation in the cytoplasm. Especially for disulphide bond formation under anaerobic conditions, periplasmic protein disulphide isomerases (PDIs) with dithiol-oxidizing function are needed. The periplasmic PDIs DsbAB (oxidizing) and DsbD (reducing) affect the maturation of c-type cytochromes (102,307,407). Another example of a periplasmic PDI is CcmG which is also involved in c-type cytochrome biosynthesis (308).

4. PROTEASES
Signal peptidases I and II are involved in the cleavage of the N-terminal signal sequences (553) but to date no proteolytic enzymes are known to be specifically involved in controlling degradation or stability of respiratory complexes depending on their stage of assembly. Other nonspecific cellular proteases such as FtsH and HtpX may just recognize unfolded or misfolded polypeptides. FtsH, otherwise known as HflB, degrades SecY (11) and subunit a of the F-Fo-ATP synthase (10) when they are not complexed. Also, other proteins tagged with the SsrA degradation peptide are substrates for FtsH (201). FtsH is found in complex
with HflC and K which are thought to modulate its proteolytic activity (249). These proteins copurify with YidC suggesting protein quality control at the very early stages of protein insertion (493). FtsH degrades membrane proteins by first dislocating the substrate from the membrane and subsequently degrading it. Reconstitution of this process in proteoliposomes showed that FtsH alone is capable of substrate dislocation (250). The metalloprotease HtpX is also involved in protein quality control and it has been suggested to work together with FtsH by introducing endoproteolytic cleavages in cytoplasmic regions of misfolded proteins which would be subsequently recognized by FtsH (12). In E. coli, cell envelope stresses are dealt with by the two-component system CpxR-CpxA. The Cpx system regulates HtpX and in doing so can combat envelope stress caused by misfolded membrane proteins (440).

There needs to be a careful coordination of cofactor biosynthesis, cofactor insertion and subunit recruitment both prior to and following protein targeting and attachment to the cytoplasmic membrane. Coordination data is, however, scarce. In the next section, the membrane biogenesis of individual respiratory complexes of E. coli is discussed. Where applicable cofactor insertion, assembly in the cytoplasm, targeting to the membrane, insertion into or translocation across the membrane and protein complex assembly following membrane insertion are handled. Lastly, we address the hypothesis that respiratory complexes are organized to specialized locations in the cytoplasmic membrane known as “respirazones”.

5. NADH DEHYDROGENASES

5.1 NADH DEHYDROGENASE I

The NADH:ubiquinone oxidoreductase (NADH dehydrogenase I or complex I, abbreviated as NDH-I) is the entry point for electrons in the respiratory chain. NDH-I catalyzes the first step of electron transport by the oxidation of NADH, thus providing two electrons for the reduction of quinone. Electron transfer from NADH to ubiquinone is coupled with the translocation of protons across the membrane thus contributing to the generation of a proton motive force (PMF) (74,347,515,536). NDH-I is the preferred NADH dehydrogenase under anaerobic growth conditions (475). Unlike the NADH dehydrogenase II, the proton-pumping NDH-I is energy conserving, which is needed under the “energy limited” conditions of growth without oxygen.
Chapter 1

All NDH-I subunits are homologues of the subunits of complex I core enzyme in mitochondria and the bacterial complex represents a structural minimal form of a proton-translocating NADH:ubiquinone oxidoreductase (271). NDH-I is one of the largest protein complexes in the bacterial membrane with a molecular mass of 550 kDa. It is a L-shaped complex composed of 13 polypeptides, named NuoA through to N where NuoC and D are fused proteins (27,419). The complex contains one FMN cofactor and 9 [Fe-S] clusters (158,536,537). The “peripheral arm” comprises of 6 subunits, namely NuoB, C/D, E, F, G and I. It extends into the cytoplasm and contains all the redox groups (420,461,540). Subunits B, C/D and I form the connecting subcomplex of the peripheral arm while the membrane part of NDH-I consists of the other 7 subunits namely NuoA, H, J, K, L, M and N (119,234,235). The membrane arm plays an important role in proton translocation and quinone binding (186,236) but little is known about the exact function of each subunit (160,209). The arrangement of both arms has been visualized by EM (28,159,190). NuoA, J and K may directly interact with the peripheral arm while NuoM and L are distantly located from the peripheral arm (28,119,237).

Assembly of complex I has been studied extensively in eukaryotes. In Neurospora crassa, the peripheral and membrane arms assemble independently and fuse en bloc to build the holoenzyme (332). The membrane arm in turn is assembled from two intermediates with the participation of two chaperones (262,508). With the human complex I, assembly is even more complex (18,486,507). In bacteria, phylogenetic analysis has revealed that NDH-I evolved from preexisting modules for electron transfer and proton translocation (157,160,161,296). The soluble NADH module comprises of an electron input module which is connected to the hydrophilic part of the amphipathic dehydrogenase module (154). The hydrophobic part of the dehydrogenase is connected to the “transporter” module containing homologues of monovalent cation/proton antiporters (157,296,459). The assembly intermediates may resemble these modules (69,270). Isolation of E. coli NDH-I in the detergent dodecylmaltoside (DDM) at pH values above 7.5 resulted in the separation of the complex into 3 parts – the NADH dehydrogenase, connecting and membrane or proton translocating domains (270) – supporting the idea that the complex is build from subcomplexes. The NADH dehydrogenase subcomplex, comprised of subunits E, F and G, was found to be fairly stable and capable of electron transport while the other 2 parts disintegrated upon further purification (270). Furthermore the FMN and [Fe-S] clusters were identified in the dehydrogenase part.

Further evidence that NDH-I assembly intermediates resemble functional modules
Introduction
came from recent studies using *E. coli* strains from the “Keio collection” of the Nara Institute of Science and Technology, Japan (23) in which the genes contained in the *nuo* operon were disrupted by a kanamycin resistance cartridge. The NADH dehydrogenase subdomain was found to form in the absence of *nuoC/D,H,I,J,K* and *N* (428), and was recovered from the cytoplasm containing the [Fe-S] clusters N1a and N1b. There were no other subunits found associated with the NuoEFG subcomplex, indicating that it likely represents an assembly intermediate. Deletion of any of the *nuo* genes resulted in a complete loss of membrane associated, electron transport chain coupled activity but the location and assembly of the remaining subunits differed between the mutant strains (428). In mutant strains where *nuol* was disrupted, a fully assembled complex was found in the membrane with only NuoL absent but without detectable NADH oxidase activity. This complex was unstable and the composition is therefore unknown but likely contained all subunits except NuoL (28,209). This could indicate that NuoL is assembled with the other NDH-I subunits in a late stage of the assembly process. In a proteomic study on GroEL/ES, NuoC/D, E and F were found encapsulated by GroEL/ES suggesting that they are folded together prior to association with the other subunits of NDH-I (245).

Oxa1-related proteins may have a role in complex I biogenesis. In *Neurospora crassa* the absence of Oxa1 destabilizes 29.9 and 24 kDa protein subunits (331). Similarly in *E. coli*, YidC depletion under anaerobic conditions led to a decrease in the amount of NuoK in the membrane as well as a severe growth defect. (376).

5.2 NADH DEHYDROGENASE II
NADH dehydrogenase II (NDH-2) is a single protein that catalyzes the transfer of electrons from NADH to ubiquinone but unlike complex I, does not couple the reaction to proton transfer. It contains the flavin cofactor FAD as its sole redox centre (222). NDH-2 is the dominant entry point for electrons into the respiratory chain under aerobic conditions. It is inserted into the membrane via the SecYEG translocase and has been shown to have no dependence on YidC for insertion (502).

6. SUCCINATE DEHYDROGENASE AND FUMARATE REDUCTASE
*E. coli* cells are able to reversibly oxidize succinate to fumarate, reactions that were first thought to be catalyzed by a single enzyme. Classical genetic approaches, however, revealed that *E. coli* possesses two enzymes, succinate dehydrogenase
and fumarate reductase (208). Both enzymes catalyze succinate oxidation but the aerobically-expressed succinate dehydrogenase has a succinate-oxidizing to fumarate-reduction activity ratio of 25:1 while the anaerobically-expressed fumarate reductase a ratio of 1:1.5 (208). The enzymes share an αβγ structure where the catalytic subunits (αβ) are cytoplasmically located and anchored to the membrane by two integral membrane proteins (γ). Although these homologues share many catalytic and structural properties with each other the assembly of the two enzymes is not identical.

6.1 SUCCINATE DEHYDROGENASE

Succinate dehydrogenase/succinate-ubiquinone oxidoreductase (SQR) is also known as complex II because of its similarity with the protein complex found in mitochondria. It is a tricarboxylic acid (TCA) cycle enzyme associated with the membrane of *E. coli* (4,198), catalyzing electron transfer from succinate to ubiquinone with the formation of fumarate. SQR does not constitute an proton translocating site in the respiratory chain and it is primarily expressed under aerobic growth conditions. It is composed of 4 non-identical subunits, encoded by the *sdhCDAB* operon, and 5 prosthetic groups (98,252,529,539).

The catalytic subunits are responsible for the reduction of succinate and are encoded by *sdhA* (flavoprotein) and *sdhB* (Fe-S protein). SdhA is the largest hydrophilic subunit. It contains covalently bound FAD (98) while the other hydrophilic subunit, SdhB, contains 3 different [Fe-S] clusters namely S-1 [2Fe-2S], S-2 [4Fe-4S] and S-3 [3Fe-4S] (329,529). SdhA and B are located on the cytoplasmic side of the membrane and are bound to the membrane integral proteins SdhC and D. The cofactor heme b bridges the two subunits in the membrane to form cytochrome b₅₅₆. The heme axial ligands are formed by His-84 of SdhC and His-71 of SdhD (505). Once the intact cyt b₅₅₆ is present in the membrane, the soluble catalytic portion binds to the membrane. (252,322). If either SdhC or D is absent there is unstable succinate dehydrogenase activity in the cytoplasm and no succinate-ubiquinone oxidoreductase activity (329). Also, the ratio of SdhA to SdhB is perturbed in the absence of SdhCD (10 Fold higher than in the wild type). If SdhC and D are added, SdhA and B are found at the membrane with succinate dehydrogenase and succinate-ubiquinone oxidoreductase activity which is stable and fully retained (329).

Although the genes are ordered in the *sdh* operon, the assembly of SQR does not require cotranslation of the subunits. When *sdhC* and *sdhDAB* are expressed from
Introduction

Different plasmids in *E. coli*, growth on succinate was supported (329). This in contrast to the structurally similar fumarate reductase (see below). The heme cofactor plays an interesting role in the post-translational assembly of SQR. Tran *et al.* constructed an *E. coli* mutant strain which was completely heme free (476). In this strain, SQR assembles correctly and retains its physiological function but is sensitive to low concentrations of detergents indicating some loss of the structural integrity (476). Also, the succinate oxidase activity observed in the heme-free mutant is not stable over time. When the SdhAB dimer is removed from SQR, the enzyme is also sensitive to oxygen and temperature (290). Taken together, the absence of heme may affect the stability of SdhCD in the membrane and/or the attachment of the catalytic subunits SdhAB. Fumarate reductase does not contain heme in its membrane domain although the general architecture of the membrane anchors is the same (217, 290).

6.2 FUMARATE REDUCTASE

Fumarate reductase/menaquinol-fumarate oxidoreductase (QFR) is expressed under aerobic conditions but at a low level (193). Its main role is under anaerobic growth conditions when fumarate is used as a terminal electron acceptor. The crystal structure revealed a butterfly-shaped dimer with an extensive buried surface between the two complexes (217). The dimer is however unlikely to have a mechanistic role since the distance between the redox centers of the electron transfer chain of the two complexes is too large. The genes are encoded in the order *frdABCD* but FrdC and D must insert into the membrane before FrdA and B can attach (268). FrdAB are both essential for catalytic activity (267) and contain FAD and three [Fe-S] clusters, respectively. A soluble catalytic FrdAB dimer can be isolated but the attachment of these subunits to the membrane anchor subunits has been shown to occur as separate events, probably following cofactor insertion in the cytoplasm (268) (Figure 5).

The assembly of the subunits to form an active fumarate reductase requires cotranslation of the subunits FrdC and D indicating that the subunits interact during or directly after translation (99). If the genes encoding these subunits are present on separate plasmids, the cells are unable to grow on medium supplemented with glycerol and fumarate. The enzyme complex that is assembled is, however, membrane located and can catalyze the reduction of fumarate by reduced benzyl viologen but is not able to interact with quinone analogues and is thus functionally uncoupled to the rest of the respiratory chain. Neither FrdC or D
alone are able to bind FrdA and B to the membrane resulting in no transfer of reducing equivalents to ubiquinone (99). This highlights that although succinate dehydrogenase and fumarate reductase are structurally very similar and can even catalyze the same reactions, their assembly is not identical.

Very little is known about the insertion requirements of the membrane anchor subunits. It was shown recently that under YidC-depleting conditions cells no longer grow on fumarate and glycerol and that the coupled activity of fumarate reductase was reduced (376). Also, the amount of FrdAB associated with the membrane was reduced suggesting a role for YidC in the biogenesis of fumarate reductase, most likely in the insertion of FrdC and/or D.

Taken together, a model for fumarate reductase assembly can be proposed in which FrdCD insert into the membrane via the Sec pathway with the assistance of YidC in a manner that is dependent on cotranslation of the subunits (376,403). FrdB binds its [Fe-S] clusters and then subsequently binds to the membrane located FrdCD dimer. Only then, would the FAD-containing FrdA bind to the FrdBCD complex as illustrated in Figure 5 (267,268).

FIGURE 5. Assembly of fumarate reductase. FrdC and D are inserted into the membrane where upon they form a dimer. FrdC and D must be synthesized cotranslationally for a membrane-bound enzyme complex to form. FrdB binds 3 [Fe-S] clusters and then subsequently binds to the FrdCD membrane anchor. After binding its FAD cofactor, FrdA assembles with FrdBCD to form an active enzyme complex (PDB# 1FUM). FrdA,B,C and D are coloured light to dark grey, respectively.
E. coli encodes two separate formate dehydrogenases, namely formate dehydrogenase N (FDH-N) and dehydrogenase O (FDH-O) which are expressed under anaerobic and aerobic conditions respectively (175,488). Both the enzymes are membrane-bound heterotrimers consisting of two cofactor-binding peripheral membrane subunits, i.e., α and β which associate with a third integral membrane γ subunit.

7.1 FORMATE DEHYDROGENASE N (FORMATE DHN, FdnGHI)
Under anaerobic conditions in the presence of nitrate, FDH-N is induced. FDH-N is encoded by the fdnGHI operon (37). It catalyzes the oxidation of formate to carbon dioxide and H⁺. It is a component of the nitrate respiratory pathway and under anaerobic conditions formate oxidation is coupled to nitrate reduction via lipid-soluble quinone. The α subunit, encoded by fdnG, is a membrane peripheral selenomolybdo-protein that associates with the β subunit. The β subunit which contains four [4Fe-4S] clusters is encoded by fdnH. It contains a C-terminal TMS which anchors the αβ subcomplex to the inner membrane of E. coli. The γ subunit, encoded by fdnI, is an integral membrane protein containing 4 TMSs and is a b-type cytochrome (140). The high resolution structure of FDH-N shows a homotrimeric complex with the α and β subunits periplasmically located (231). The positioning of the quinones in the structure, i.e. on opposite sides of the membrane, suggests that FDH-N is involved in PMF generation which is important in anaerobic respiration. The monomer is the functional unit but trimerization may increase the stability of the enzyme complex. FdnGH trimers show very extensive protein-protein contacts with each other in the mature enzyme therefore there must be a mechanism to prevent FdnGH “monomers” from interacting before transport. This could be chaperone mediated.

The structure is missing the N-terminal 33 amino acids of FdnG which contain the Tat signal sequence (38,356,415) essential for FDH-N assembly (58,415). The Tat signal peptide is only present on the FdnG subunit, and yet both FdnG and H are transported across the cytoplasmic membrane (447). The Tat system must therefore transport FdnH to the periplasm following complex formation with FdnG implying that the targeting of FdnG to the Tat translocon must be prevented until complex formation with FdnH occurs. This function may be provided by a “cofactor chaperone” possibly formed by formate biogenesis factors FdhD and E.
Chapter 1

(292). FdhE is only required for the Tat-dependent biosynthesis of formate dehydrogenases and not for the assembly of non-exported formate dehydrogenases. In addition to the accessory proteins FdhD and FdhE, FdhF is also required for FDH-N biogenesis (292,427,453).

Sequence analysis of probable Tat-dependent substrates suggested that five of them contained C-terminal TMSs. This included FdnH which contains one TMS at the C-terminus with an N_out-C_in topology (196,231). FdnH was shown to be targeted to and translocated across the membrane via the Tat system and the insertion of the C-terminal α-helix to occur independently of the Sec translocon and YidC (196). A likely mechanism is that the C-terminal tail anchor behaves as a “stop transfer” domain. The bulk of the protein would be transported to the periplasm by the Tat translocon. The hydrophobic C-terminal helix would however not be completely translocated by the Tat translocon and would escape laterally into the lipid bilayer. YidC is not involved in the lateral release of the TMS and it is unclear what clears the Tat translocation channel or what integrates the TMS. Alternatively, Tat-dependent TMS insertion could occur by a “periplasmic re-entry” mechanism. After export of the entire enzyme the C-terminal TMS would then associate with the inner membrane spontaneously or via another protein, possibly FdnI, from the periplasmic side. The membrane biogenesis of the γ subunit, FdnI, has not been extensively studied and has been postulated to proceed via the Sec-pathway (196).

7.2 FORMATE DEHYDROGENASE O (FORMATE DO, FdoGHI)

The E. coli genome encodes a second FDH, the FDH-O which is encoded by the fdoGHI operon. FDH-O catalyzes the oxidation of formate to carbon dioxide and H⁺ in a NAD-dependent reaction. FdoH, like FdnH, has a C-terminal TMS that is inserted into the membrane via the Tat translocon (196). FdoH has an N_out-C_in topology suggesting that FdoGH are targeted as a preformed dimer but not translocated across the cytoplasmic membrane by the twin arginine leader present on FdoH (36,414). Gene-fusion reporter studies have shown that FDH-O is cytoplasmically orientated (36). If this holds true, this enzyme is an interesting example of protein complexes that use the Tat system for targeting to but not translocation across the membrane.
Introduction

8. HYDROGENASES

The *E. coli* genome encodes 3 hydrogen-evolving hydrogenases, namely hydrogenase 1, 2 and 3, which catalyze the production and consumption of hydrogen gas (5). Hydrogenase 1 and 2 are also often referred to as [NiFe] hydrogenase 1 and [NiFe] hydrogenase-2 isoenzyme, respectively. The protein complexes are encoded by the operons *hyaA-F*, *hybO,A-G* and *hycA-I* (60,305,306,413). All the structural components of the hydrogenases are contained in the operon together with proteins essential for regulation and maturation of the enzyme complexes. A fourth operon, *hypA-F*, encodes proteins involved in the maturation of all 3 hydrogenases (284). Although high resolution structures are available for hydrogenases from sulfate-reducing organisms, for example *Desulfuvibrio gigas* (509), these belong to a different subfamily of hydrogenases (532) and the structure of any of the *E. coli* enzymes cannot be inferred from them.

8.1 HYDROGENASE 1 (*hyaA-F*)

Hydrogenase 1 consists of 3 proteins: the small chain HyaA, the large chain HyaB, and an integral membrane protein HyaC that is probably the *b*-type cytochrome subunit. HyaA contains a C-terminal TMS that is inserted into the membrane via the Tat translocon (196,414). In addition to the structural genes, the maturation proteins HyaDEF are contained in *hya* operon. HyaD encodes for a protease possibly involved in the C-terminal processing of HypB (inferred by sequence similarity) while the functions of HyaE and F are not known.

8.2 HYDROGENASE 2 (*hybO-G*)

Hydrogenase 2 is encoded by one operon which contains many proteins that are not structural components of the protein complex. The large subunit contains the Ni-Fe-CO-2CN cofactors and is encoded by *hybC*. *hybO* encodes for the small subunit. It is a 39 kDa Fe-S protein with a large cleaved N-terminal Tat signal sequence of 4 kDa (413). HybO contains one TMS which anchors the HybOC subcomplex to the membrane (219) and this C-terminal TMS of HybO is solely dependent on the Tat translocon for insertion (196). HybOC form a subcomplex in the cytoplasm which is dependent on the Tat translocon for membrane targeting and translocation (390,411). The Tat signal is contained on only the small subunit and HybC is cotranslocated by the Tat translocon together with HybO by means of a “piggy-back” mechanism. The integral membrane subunit is encoded by *hybB* which is a probable *b*-type cytochrome and contains 11 potential TMSs. HybA was
previously thought to be the small subunit of the hydrogenase 2 (305) but this proved to be incorrect (413). HybA is an [Fe-S] cluster-containing protein and probably serves as the electron acceptor from the small subunit. The Fe-S and Ni-Fe-CO-2CN cofactors must be inserted into the HybO and HybC proteins, respectively before the proteins can be targeted to the membrane. This occurs in the cytoplasm and most likely at a stage prior to the association between HybO and HybC (Figure 6). HybC has been found to copurify with another protein product of the hyb operon, HybG (127,220). In the assembly of the metallocentre of HybC, HypE together with HypF prepare cyanide ligands from carbamoyl phosphate (53,359,383) which are then delivered by HypE to HybG which delivers the Fe centre to HybC. The association of HypE, together with HypD, with HybG is required for the stable association of HybG with HybC (80). HybE also functions as a dedicated Tat dependent chaperone in the biogenesis of hydrogenase 2 (219). It has been suggested that HybE masks the Tat signal peptide while the cofactors are inserted and subunits are recruited (127). In the absence of HybE, HybO is still membrane targeted while HybC remains in the cytoplasm indicating that the primary role of HybE is in proofreading rather than targeting to the Tat translocon (127). While Tat signal sequences of molybdoenzymes contain the characteristic twin arginine motif, they are unique to the enzyme. If the Tat signal sequence of TorA is replaced for that of HybO, HybO is still targeted to the membrane but HybC is not. HybE therefore specifically recognizes the signal sequence of HybO and allows the coordination of cofactor insertion and subunit recruitment so that a preformed HybOC complex can be presented to the Tat translocon for transport across the membrane. HybOC then associates with the integral membrane subunit HybB and the Tat-dependent ferredoxin HybA (196,413).

A model for chaperone-mediated proofreading has been proposed and is illustrated in Figure 6 (219). In this model the apoprotein bearing the Tat signal peptide is released from the ribosome and to prevent premature export, a proofreading chaperone binds to the Tat signal peptide. In some cases, specific or general chaperones may bind to the mature part of the protein at this stage. Cofactor insertion in the mature portion of the protein and the binding of partner subunits can then take place. All chaperones would then be released allowing the precursor to be targeted to the Tat translocon. The protein would then be translocated across the cytoplasmic membrane by the Tat machinery, the signal peptide cleaved and the mature protein released into the periplasm (219). The action of the chaperones ensures that only fully assembled substrates are presented
FIGURE 6. **Chaperone-mediated maturation of hydrogenase 2.** In the cytoplasm, HybE binds to the Tat signal peptide of HybO preventing premature targeting to the Tat translocase and mediating cofactor insertion. HybDEFG assemble the metallocentre at HybC. Once HybO and HybC have acquired their respective cofactors they bind to form a dimer. The HybOC dimer is translocated across the membrane via the Tat translocon and is reliant on the Tat signal peptide contained in HybO. HybO contains a C-terminal TMS which is inserted into the membrane by the Tat translocon. Concomitantly, the Tat-dependent ferrodoxin HybA acquires its [Fe-S] clusters and is translocated to the periplasm. HybB contains a probable $b$-type cytochrome and is most likely integrated into the membrane via the Sec translocase whereupon it undergoes maturation. It is only then that the holoenzyme complex is formed. Interactions between the subunits of hydrogenase 2 in the holoenzyme are purely illustrative.

This would increase the efficiency of Tat transport as opposed to quality control being performed by the Tat translocon itself. Abortive engagement would be prevented and the competition with substrates not ready for export reduced.

**8.3 HYDROGENASE 3**

Formate hydrogenlyase (FHL) comprises of formate dehydrogenase-H, unidentified electron carriers and hydrogenase 3. It catalyzes a non-energy conserving pathway in which molecular hydrogen and carbon dioxide are released.
from formate. The genes are encoded by the \textit{hyc} operon which encodes structural proteins as well as those involved in regulation and maturation of the hydrogenase (60). HycA is a regulatory protein (417). HycBCDEFG are all structural components of the FHL. HycB, E, F and G all contain [4Fe-4S] clusters while HycE also contains the nickel cofactors. HycF is a probable electron transfer protein for hydrogenase 3. HycC and D are integral membrane proteins containing 12 and 8 TMSs, respectively (60,417). Subunits HycE, F and G are similar to the connector module of complex I while HycC and D share similarity with the hydrophobic membrane subcomplex of NDH-I. The assembly of the FHL complex may therefore resemble these subcomplexes as has been postulated for NDH-I. HycH and the protease HycI are maturation proteins required for processing the C-terminal of HycE (393).

In addition to the proteins encoded by the \textit{hypA-G} operon, the general chaperones SlyD and GroEL/ES are required for hydrogenase 3 assembly (80,548).

9. TMAO AND DMSO REDUCTASES

In the absence of alternative substrates, \textit{E. coli} can use trimethylamine N-oxide (TMAO) and dimethyl sulfoxide (DMSO) as terminal electron acceptors. These compounds are widespread in marine environments.

9.1 TMAO REDUCTASE

TMAO reductase encoded by the \textit{torCAD} operon, is able to reduce TMAO to trimethylamine (TMA) under anaerobic conditions. The enzyme has two periplasmically located subunits, TorA and TorC (103). TorA is a large molybdoprotein and TorC is a \(c\)-type cytochrome anchored to the cytoplasmic membrane by an N-terminal TMS (216). In the absence of TorA, TorC is still present in the membrane and contains its heme cofactor (304). TorC is likely anchored in the membrane via the Sec translocase while TorA is transported by the Tat translocon as a fully assembled protein. There is no default targeting to the Sec translocase when Tat is not present (196). Interestingly, it was found that cofactor insertion rather than the signal sequence is critical for transport (219).

The biosynthesis of TorA is governed by TorD which is a member of a chaperone family dedicated to molybdo-protein assembly (483). There is a wealth of literature describing the involvement of TorD in cofactor loading (171-174,214,215,367). TorD interacts with mature regions of heat-denatured but not native TorA which lead to the proposal that the TorD family of molecular chaperones functions in the insertion of molybdopterin cofactors into the apoprotein (367). TorD also performs
Introduction

a Tat proofreading function, preventing premature targeting of TorA until all biosynthetic processes are complete (219). The Tat proofreading function of TorD involves a direct specific interaction with the entire TorA signal peptide and not just the Tat motif sequence (196,219). Tat targeting and the Tat signal peptide required for recognition and binding by TorD are distinct and the two processes may not be mutually exclusive. TorD has been proposed to bind TorA at two binding sites – one for the Tat signal sequence and one for the “mature” part of the protein. TorD from S. massilia is a dimer (477) and the monomer consists of N and C domains and a hinge region. The N- and C-domains are involved in cofactor insertion as indicated by site-directed mutagenesis while the hinge region is important for TorA signal recognition (219). If other chaperones are deleted, both the N- and C-domains are needed for TMAO activity (219). TorD has also been implicated in preventing the degradation of the Tat signal peptide (173) and stabilizing apoTorA when cofactors are limiting (174). If the signal peptide is deleted, MGD insertion still occurs reiterating that TorD recognizes more than just the signal peptide (215,464). In a TorD null mutant, a significant decrease but not an abolishment of TMAO reductase activity in the periplasm was observed. Although the N-terminus of TorA is unique, replacement of the TorA signal sequence by that of HybO had no effect on the TMAO activity in the periplasm. This indicates that the primary role of TorD must be molybdenum cofactor insertion. This is contrast to HybE where the Tat proofreading function is most crucial (219). TorD is homologous to DmsD and NarJ which are required for the biogenesis of the DMSO and nitrate reductases, respectively.

9.2 DMSO REDUCTASE
The DMSO reductase is encoded by the dmsABC operon which is expressed under anaerobic conditions in the absence of nitrate (44,46). The DMSO reductase can also reduce TMAO (522) but has a higher affinity for DMSO (45). The dms operon is uniquely required for growth on DMSO, a function which cannot be filled by the TMAO reductase. DMSO reductase consists of 3 subunits all of which are essential for growth in DMSO and membrane localization of the enzyme complex (299,404). The catalytic dimer, formed by DmsAB, is anchored to the cytoplasmic side of the membrane by the anchor subunit DmsC. The interaction between the DmsAB dimer and DmsC is mediated by the C-terminal region of DmsB since the deletion of the last 30 C-terminal amino acids of DmsB results in cytoplasmically located
Chapter 1

DMSO reductases activity (406). DmsC not only functions to anchor the catalytic dimer to the membrane but is also involved in the physiological transfer reaction from menaquinone and increases the stability of the reductase activity. If DmsA, which contains the MGD cofactor, is absent, DmsB, the Fe-S cluster containing subunit, is rapidly degraded (82,406).

DmsA has an N-terminal twin arginine leader which functions as a membrane targeting signal and is essential for the stability of the holoenzyme (405). The Tat system catalyzes the targeting of the DmsAB dimer to the cytoplasmic face of the membrane but does not translocate the dimer across the membrane. Without the Tat system, DmsAB subunits accumulate in the cytoplasm and the reductase activity that is present is very labile suggesting rapid proteolytic degradation of the mistargeted DmsAB dimer. This indicates that a functional Tat system may be required for processing of DmsA to its mature form (403,415). In *E. coli* strains devoid of the Tat system, there is no growth in either glycerol and DMSO or TMAO-containing medium (403,482). Also, in the absence of membrane-targeted DmsAB, DmsC accumulates in the membrane (482) indicating that the biogenesis of DmsC is a separate event and is not dictated by the maturation state of DmsAB.

The leader sequence of DmsA does not simply function as a stop-transfer signal as its replacement for the TorA signal sequence did not result in translocation (405). There must therefore be additional information in either the leader sequence of DmsA that is not present in that of TorA or in the mature DmsAB polypeptides that prevents translocation. A similar phenomenon has been observed for FdoGH (36). This shows that both the twin arginine leader sequence and the mature polypeptide(s) contribute to the final localization of the enzyme. Like the TMAO reductase, the DMSO reductase has its own dedicated chaperone, DmsD, which is part of an operon encoding an apparently cryptic DMSO reductase homologue (352). DmsD interacts with the Tat signal peptide of DmsA (352) and the Tat translocon (357,381). DmsD contains specific binding sites for both the signal peptide and the cofactor-less unfolded mature portion of DmsA (352,367). Unexpectedly it is able to interact specifically with the precursor forms of both DmsA and TorA even though these proteins have no obvious sequence similarity beyond the twin arginine motif (512).

In the absence of DmsD, no cellular DMSO activity is observed (352) while in strains devoid of Tat components, only a mislocalization of the activity to the cytosol occurs (59,415,416). This led to the proposal of the “cofactor chaperone”
Introduction

model (42,411). This model suggests that the signal peptide of Tat precursors is prevented from interacting with the Tat translocon until cofactor insertion into the precursor protein is completed. In addition, DmsD translocation may also occur via an “escort chaperone” mechanism since if the signal peptide of DmsA is fused to cytochrome c, the fusion protein is exported via the Tat system (410).

10. NITRATE REDUCTASES

*E. coli* encodes for three nitrate reductases – two homologous NAR nitrate reductases which are located on the cytoplasmic side of the membrane and NAP which is a periplasmically located nitrate reductase (50,65,97,175). These complexes catalyze the reduction of nitrate to nitrite under anaerobic conditions. Nitrate is the preferred terminal electron acceptor in the absence of oxygen which is reflected in the transcriptional repression of other terminal reductases by nitrate (194,488).

10.1 NAR NITRATE REDUCTASE

The NAR nitrate reductases are contained in the operons *narGHIJ* and *narZYWV*. Nitrate reductase A (NarGHI) has been studied extensively and will be discussed in detail. When *E. coli* is grown anaerobically in the presence of high levels of nitrate, the respiratory chain that is formed is terminated by NarGHI otherwise known as the quinol nitrate oxidoreductase (50). The heterotrimeric NarGHI has an αβγ structure encoded by *narGHI* respectively. The α and β subunits can be purified as a soluble NarGH dimer in the cytoplasm (24,283) showing that these 2 subunits associate to form a stable subcomplex before association with the membrane γ subunit (Figure 7). The NarGH dimer is capable of nitrate reduction but without the membrane integral γ subunit this reaction is not coupled to the respiratory chain. NarG contains an Mo-bis-MGD cofactor and [Fe-S] cluster and NarH contains 4 [Fe-S] clusters (43,49,395). The NarGH subcomplex is anchored to the cytoplasmic side of the membrane by NarI. NarI contains 5 TMS and two b-type hemes (40,117). It is a quinol oxidase and transfers electrons from ubiquinol to NarH. The high resolution structure of nitrate reductase revealed a dimer and importantly showed it to be cytoplasmically located as suggested by earlier studies (153,229). This revealed a mechanism by which electron transfer from formate to nitrate could be coupled to the generation of a PMF.

Contained in the same operon is the assembly factor NarJ without which the nitrate reductase complex cannot assemble (51,128). If *narJ* is deleted, there is no formate-linked membrane-bound enzyme activity but there is soluble nitrate reductase activity.
reductase activity in the cytoplasm (51, 128). If NarJ is absent, NarI is unaffected but there are low levels of the $\alpha\beta$ subcomplex in the cytoplasm with low activity. The NarGH peptides are partially degraded indicating an accumulation of immature dimer which is susceptible to proteolysis (51, 128, 444). Based on sequence similarity NarJ belongs to the TorD family of molecular chaperones. The chaperone binds two distinct sites on apoNarGH and binding to the one site allows interaction of apoNarGH with the molybdenum cofactor biosynthesis machinery (503). The second binding site includes the first 40 amino acids of NarG which are responsible for the interaction of NarG with NarI (43). It is possible therefore, that NarJ functions primarily to prevent membrane anchoring of the apoenzyme before molybdenum insertion is complete as well as maintaining the protein in a soluble form during the insertion of the cofactor. In the absence of NarJ, NarGH interacts prematurely with NarI forming an inactive complex (504). A role for the Tat machinery in the targeting of the NarGH dimer to the membrane remains controversial. It has been suggested that although NarG contains no cleavage site or typical Tat signature a vestige twin arginine motif exists (483). NarJ has been shown to bind the first 15 amino acids of NarG and in a $\Delta$tatA-E strain less is found associated with the inner membrane compared to the wild type (88). However, in another study it was reported that $\Delta$tat strains were able to grow under anaerobic conditions when nitrate was present indicating that NarGH do not utilize the Tat system in agreement with the absence of a canonical Tat signal sequence (403, 482). Also after membrane association, NarG is not cleaved which usually occurs with Tat substrates. If the N-terminus of NarG is absent, NarJ still binds and coordinates cofactor insertion although the molybdenum associated with NarG is significantly reduced while the [4Fe-S] remains the same (265) The NarGHI complex still forms but NarI associated with this complex contains only the distal $b$-type cytochrome ($b_D$) and not the proximal one ($b_P$). This indicates that the $b_P$ is inserted late in the maturation process and also, since in the absence of NarGH and J NarI is still properly matured, that NarJ ensures the proper timing of membrane anchoring (287, 394). The $b$-hemes of NarI are sequentially added and this may also be the case for other respiratory complexes such as succinate dehydrogenase.

NarJ-mediated assembly of nitrate reductase A is illustrated in Figure 7. In summary, NarJ interacts with apoNarGH early on, promoting the insertion of the [4Fe-S] clusters and molybdenum in a sequential manner. In parallel, NarI receives two hemes without the aid of NarJ, with the $b_D$ being inserted before the $b_P$. The mature NarGH complex is directed to the inner membrane for interaction
FIGURE 7. *NarJ*-mediated biogenesis of *NarGHI*. First, the [Fe-S] clusters are added to NarH whereupon NarJ binds to NarG. NarJ mediates [Fe-S] cluster and Mo-bisMGD cofactor insertion and prevents premature targeting of the NarGH dimer to the membrane. In parallel, NarI is inserted via the Sec pathway into the membrane whereupon maturation takes place. The $b$-type hemes are sequentially inserted. NarJ dissociates from the NarGH complex allowing targeting to the membrane and anchoring of the NarGH dimer to the matured NarI to form NarGHI (PDB# 1Q16). NarG, H and I are coloured medium, light and dark grey respectively. Rounded squares indicate mature subunits.

with the mature NarI subunit. Whether this involves the Tat translocon is uncertain.

10.2 PERIPLASMIC NITRATE REDUCTASE (NAP)

The periplasmically-located nitrate reductase (NAP) is contained in the operon napABCDFGH (192). The structural components are encoded by napABC (224). NapA contains the molybdopterin cofactor and a [4Fe-4S] cluster as well as a Tat signal sequence (387). NapB is a diheme $c$-type cytochrome which contains an N-terminal signal sequences cleaved after translocation across the inner membrane. NapA and B form a soluble periplasmic enzyme. NapC, which is anchored to the inner membrane by an N-terminal hydrophobic segment, is required for *in vivo* nitrate reductase activity and transfers electrons from its tetraheme cytochrome $c$ centre to the NapAB complex (385). NapA, B, C and D are all essential for the formation of an active NAP complex while NapF, G and H are not (373). NapF is a
cytoplasmic Fe-S protein that is required for the assembly of the [Fe-S] clusters in NAR (349) and has been shown to interact with NapA (334). NapD is a Tat signal peptide-binding protein which binds directly to NapA preventing premature interaction between NapA and the Tat translocon (289). In the absence of NapD there is no periplasmic nitrate reductase activity (373). NapD has a β-α-β-β-α-β structure known as a “ferrodoxin-type” fold that is different to that of the TorD family which have an all helical fold (289). Therefore, there is structural diversity in the chaperone-mediated quality control systems of the Tat pathway.

11. NITRITE REDUCTASE

Under anaerobic and microaerobic conditions when low levels of nitrate are present, the periplasmic nitrate reductase and the periplasmic nitrite reductase (Nrf) systems are expressed (372). Also known as the cytochrome c nitrite reductase, the enzyme catalyzes nitrite reduction to ammonium in a formate-dependent reaction.

The nitrite reductase uses the electrons donated from formate, quinone and subsequently the nitrite reductase subunit NrfB. The nitrite reductase is contained in a large operon nrfABCDEFG (212). The active enzyme complex is only encoded by nrfAB. NrfA is a periplasmic pentaheme cytochrome c552 (135) and NrfB is a pentaheme c-type cytochrome. Both NrfA and NrfB form decaheme homodimers (25,95).

All the proteins encoded by the nrf operon are required for the formate-dependent reduction of nitrite but the reduction of nitrite to ammonium is catalyzed by NrfA. Also, contained in the nrf operon are NrfC, a periplasmic [4Fe-4S] cluster-containing ferredoxin, and NrfD, an integral membrane putative quinol dehydrogenase (39,368). These proteins are probably involved in the transfer of electrons from the quinone pool to the c-type cytochromes. NrfC contains a Tat signal sequence and is most likely translocated across the membrane by the Tat translocon following [Fe-S] cluster insertion in the cytoplasm. The other proteins encoded by the nrf operon are involved in biogenesis of the reductase. NrfE belongs to the CcmF family and is involved, together with NrfG, in the biogenesis of the c-type cytochromes contained in NrfA and B. NrfF belongs to the CcmH family but is not essential for nitrite reductase biogenesis. Although these proteins are homologous to other Ccm proteins, the cytochrome biogenesis proteins contained in the nrf operon appear to have evolved to perform specialized roles in the biogenesis of c-type cytochromes in the nitrite reductase enzyme only
Introduction

(191,192,212) and are not used by other c-type cytochromes. The heme ligation occurs on the periplasmic side, which has been observed for other bacterial cytochromes (355,407,467,468).

12. TERMINAL CYTOCHROME OXIDASES

12.1 CYTOCHROME OXIDASE

Cytochrome o oxidase, or cytochrome bo oxidase, is a quinol oxidase that accepts electrons from quinol and uses them for the reduction of molecular oxygen. It contains a heme-Cu cofactor (Cuₜ) and is essential when oxygen is present at high concentrations. The protein complex is contained in the cyoABCDE operon. The structural proteins are encoded by the genes cyoABCD while cyoE encodes for a membrane located farnesyl transferase that is involved in heme o synthesis (398,399). Subunit I, encoded by cyoB, serves as the catalytic center for dioxygen reduction and proton pumping. It is an integral membrane protein with 15 TMS and contains all the redox metal centers in the oxidase complex i.e. a low spin heme b (cytochrome b₅₆₃.₅), and a high-spin heme o-Cuₜ binuclear centre (478). Subunit II is encoded by cyoA and subunit III, whose function is not known, is encoded cyoC. Subunit IV, encoded by cyoD, assists in Cu incorporation into subunit I (400) but once the enzyme and cofactors are assembled, subunit IV is dispensable and active cytochrome oxidase consisting of only CyoA and CyoB has been isolated (251). Dimers of cytochrome o oxidase have been isolated (3). The dimer is, however, less abundant in vivo than the monomer and highly sensitive to detergents (450). Even though the subunits assemble in a 1:1:1:1 stoichiometry they are not translated in this stoichiometry. Subunits I and II are produced in limiting amounts.

If cyoBCDE are all deleted, CyoA is still detected in the membrane and if cyoA is deleted, CyoB is still present indicating that the subunits are not dependent on each other for membrane insertion or stability (328). CyoA has two TMSs and a large periplasmic domain. It has an N-terminal extension which is lipid modified during membrane insertion following which the signal sequence is removed by signal peptidase II (285). Membrane insertion of CyoA has been postulated to occur as a two step process in which the N-terminus insertion is mediated by YidC and that of the C-terminus by the Sec translocase (87,126,494,494). Using truncations of CyoA consisting of either the signal peptide and TMS 1 or TMS 2 and the C-terminal periplasmic domain it was shown that these truncates had differing
insertion requirements (87). Based on these observations it has been proposed that YidC mediates insertion of the signal peptide and TMS 1, inserting them as a helical hairpin, following which the Sec translocase inserts TMS 2, translocating the large C-terminal domain to the periplasm (86). The insertion process is strictly sequential unless the cytoplasmic loop is lengthened considerably allowing the C-terminal TMS to insert via the Sec translocase independently of the YidC-dependent N-terminal TMSs. CyoA does not require the PMF for insertion most likely owing to periplasmic regions containing no net charge (87).

In yeast mitochondria the cytochrome c oxidase consists of 12 subunits, with the 3 largest subunits Cox1, Cox2 and Cox3 forming the catalytic core of the enzyme. These core subunits are also present in the bacterial cytochrome oxidase complex and have been evolutionarily conserved. The activity of the cytochrome c oxidase is completely abolished in Δoxa1 mutants and assembly of the complex was found to be defective (15). In the absence of Oxa1, the mitochondrially-encoded subunits fail to insert properly into the inner membrane of the mitochondrial and accumulate in the matrix where they are eventually degraded (64,272). Oxa1 has been shown to interact with Cox1, Cox2 and Cox3 before their synthesis is completed. This indicates that these subunits are inserted into the inner membrane of the mitochondrion by Oxa1 in a cotranslational manner, (197,199,244,331). The biogenesis of Cox2 has been extensively studied. Cox2 has 2 TMSs with two hydrophilic tails in an N-out and C-out topology (479). In yeast, two members of the Oxa1/Alb3/YidC family are present, Oxa1 and Cox18, and both proteins are absolutely required for the insertion of Cox2 into the membrane and the export of its hydrophilic domains. Like CyoA, Cox2 has distinct requirements for the translocation of the different domains. Export of the N-terminal tail region and C-terminal domain of Cox2 from the mitochondrial matrix into the intermembrane space requires Oxa1 (197), while Cox18 is required for the insertion of the C-terminal region (412). Cox18p has also been implicated in the maintenance of steady state levels of cytochrome c oxidase and in the assembly of Cox2 with the other subunits (412).

Similarities in YidC, Oxa1 and Cox18 activities in the assembly of respiratory complexes are underscored by in vivo complementation studies. The YidC transmembrane domains can carry out the translocase functions necessary for respiratory chain assembly in Saccharomyces cerevisiae (375). A chimeric protein was constructed consisting of the mitochondrial targeting signal and N-terminal
domain of Oxa1 fused to the core 5 TMSs of YidC. This chimera was able to partially complement the respiratory growth phenotype of \textit{cox18} but not that of \textit{oxa1} deletion mutants. Addition of the Oxa1 ribosome binding C-terminal domain of the chimeric YidC protein allowed partial complementation in the \textit{oxa1} mutant strain but not of the \textit{cox18} deletion mutant (375). These observations not only revealed the evolutionarily conserved functions of the Oxa1/Alb3/YidC family but also that Oxa1 and Cox18 have evolved to perform specialized co- and post-translational functions respectively.

Assembly of the cytochrome \textit{o} oxidase in \textit{E. coli} has been shown to occur in a ordered step wise manner (Figure 8) (450). When all subunits are present, the first step in complex assembly involves the formation of a CyoCD subcomplex. CyoB then associates with the CyoCD subcomplex followed by subunit CyoA. Cofactor insertion is needed for the correct folding of CyoB and its interaction with the CyoCD complex. Heme \textit{b} is especially important for the correct folding of CyoB. In the absence of heme \textit{b}, CyoB may not be able to efficiently close its clam shell-like structure which encloses the other 3 cofactors. If the heme ligands are mutated,
there is a defect in Cus incorporation because of the resulting lower affinity of the oxidase for the metal cofactor (313). However, if mutations are made in either of the two Cus ligands (H333 and H334) normal amounts of heme b and o are present and only Cus is lacking (273,313). This indicates that heme insertion occurs before and/or independently of Cus binding. Insertion of heme cofactors into CyoB occurs also in the absence of CyoA and as mentioned above, both catalytic subunits of cytochrome c oxidase can insert autonomously into the membrane and persist (328).

12.2 CYTOCHROME bd OXIDASE
Under conditions of low aeration the high-affinity cytochrome bd oxidase is the preferred terminal reductase. The structural components are encoded by the genes cydAB which form a heterodimer (188,189). CydA contains a low-spin heme, cyt_b559, and is thought to be the quinol oxidation site (146,243). The heme d and high-spin heme, cyt_b595, are shared between CydA and B and form a binuclear centre where oxygen is reduced to H2O (104,207). If cydA is deleted both CydA and B are absent as are all the heme cofactors. If cydB is deleted, CydA is still present but the heme d and cyt_b595 are absent (188,333). The results from the cydA deletion study may be indicative of a disruption of cydB transcription rather than a dependence of CydB on CydA for insertion and stability in the membrane. Since CydA only contains the cyt_b559 in the absence of CydB, association of the two other heme moieties appears to require association of both subunits and must be a late step in the assembly of the enzyme (333). Heme ligation is however not a prerequisite for CydAB complex formation (146). CydC and D are encoded in the same operon as CydA and B and thought to form a heme transporter necessary for cytochrome bd oxidase biogenesis (33,176). YhcB was thought to be an additional component of the CydAB complex since it co-migrates with the CydAB complex in blue native PAGE (449). Further investigation, however, revealed that YhcB is not associated with the cytochrome bd oxidase and also not involved in its biogenesis (314).

There is a third terminal oxidase present in E. coli encoded by the genes cpxAB. It is a bd type oxidase but its function is unclear. It appears to be expressed under microaerobic conditions (109,456).
Introduction

13. F<sub>o</sub>F<sub>i</sub> ATP SYNTHASE

The membrane bound ATP synthase of bacteria has two important physiological functions: Firstly to synthesize ATP from ADP and inorganic phosphate utilizing the energy of electrochemical ion gradients, and secondly as an ATPase to generate transmembrane ion gradients at the expense of ATP. The order of genes in the ATP synthase operon is *atpIBEFHAGDC*. The genes are transcribed as a polycistronic message and translated by polysomes (77) and *in cis* assembly of the F<sub>i</sub> subunits may therefore occur (77). The ATP synthase is divided into 2 subcomplexes: a membrane-embedded F<sub>o</sub> subcomplex through which the protons flow and a peripheral F<sub>i</sub> subcomplex that carries the nucleotide binding sites.

In *E. coli*, the membrane-integral portion, F<sub>o</sub>, has 3 subunits in the stoichiometry ab<sub>2</sub>c<sub>10</sub> and translocates protons through the membrane driving rotation of the c subunit ring (149,180). The membrane-peripheral sector, F<sub>i</sub>, comprises 5 subunits in the stoichiometry αβγδε (71,116,228). The F<sub>i</sub> part extends from the membrane with subunits α and β alternately arranged in a hexamer surrounding the centrally located γ subunit (2). It houses the 3 catalytic nucleotide binding sites and couples rotation of the centrally located γε to synthesis or hydrolysis of ATP at catalytic sites found at the αβ interfaces (1,73,130,327,520). The γ and ε subunits form a stalk connecting the F<sub>i</sub> part with the ion-translocating F<sub>o</sub> (181). ATP synthesis/hydrolysis occurs in the different β subunits according to cooperative binding of ADP and inorganic phosphate at one catalytic site which is coupled to the release of ATP at a second site while the third site is empty (2,73). The structure of the F<sub>o</sub>F<sub>i</sub> ATPsynthase is illustrated in Figure 9.

Since the genes are transcribed as a polycistronic messenger, the α and δ subunits could assemble immediately with a 1:1 stoichiometry but this would not give rise to the δα<sub>3</sub> form as in the final F<sub>i</sub>. The α subunit is transcribed at a greater rate than the δ subunit (78) and therefore not all newly-synthesized α subunits would bind δ subunits. Alternatively, the N-terminal region of newly-translated α subunit could be sequestered in a way that binding to the δ subunit would be temporarily prevented, providing a more efficient assembly pathway (436). It has been proposed that before the δ and α subunits can interact, the α subunit must be incorporated with the β subunit into the αβ<sub>3</sub> hexagon, a process which may require the γ subunit as well (436). The N-terminal region of the α subunit would then be available for binding to the δ subunit and after binding of the first δ subunit, steric hindrance would prevent attachment of more δ subunits. There is also evidence that the binding of the α to the δ subunit leads to an increase in the
similarly to the wild type enzyme but had lost its oxidative phosphorylation activity (94). Reconstitution competition experiments with dimerization-impaired mutants of Fo-b indicated that the monomer has no significant affinity to other components of the F1 domain (445). Fo-b is therefore inserted into the membrane where it dimerizes before it interacts with the other subunits of the ATP synthase. Fo-b is arranged with Fo-a asymmetrically outside the ring-like subunit c oligomer (47,443,463). Fo-c contains two TMSs connected by a polar loop region which is exposed to the F1 binding site of the membrane (179,200). The interaction of the c-ring with the central stalk/stator is very strong as suggested by various cross-linking (202,429,480) and structural (455) studies. During ATP synthesis/hydrolysis Foγ and Foε rotate by 120º for each ATP synthesized or hydrolyzed (541) with rotation coupled by deprotonation/protonation of about 3.3 Fo-c subunits. In a complete cycle, 3 ATP molecules are synthesized/hydrolyzed and 10 protons translocated.

Fo-c over-expressed in E. coli and purified in non-ionic detergent solution self assembles into annular structures in the absence of other subunits of the ATP synthase, suggesting the ability of Fo-c to form rings is determined by the primary structure (19). Interactions of Fo-c with the other subunits of the Fo domain may also affect c-ring formation. Fo-c purified in organic solvents as a monomer can be reconstituted into liposomes in an active form in the presence of Fo-a and Fo-b (120). This also suggests that the Fo portion of the ATP synthase is capable of assembling without the assistance of chaperones or subunits of the F1 portion. Investigation of the insertion of Fo-c oligomerization mutants showed that subunit insertion and oligomerization occur in 2 distinct steps. Insertion is mediated by YidC only and is not reliant on subsequent oligomerization of the subunits (258,499,499). While the F1:Fo ATP synthase consists of 8 subunits, a ninth gene, uncI (also atpI), is encoded in the same operon. This gene encodes for a 14 kDa membrane protein (79,121) which has been implicated in c-ring assembly. In vitro experiments with UncI and Fo-c from Propionigenium modestum showed that for this bacterium, UncI interacts post-translationally with Fo-c and is essential for the formation of the c-ring (354). In E. coli however, deletion of uncI does not affect ATP synthase activity or assembly (169) and in this bacterium, YidC may be sufficient for the membrane biogenesis of the c-ring. It is worth noting that in Bacillus pseudofirmus OF4, AtpI together with AtpZ has been assigned a completely different function in magnesium uptake (206). AtpI genes from different bacteria do not exhibit strong sequence similarity and annotation according to their position in the atp operon may have resulted in
Introduction

functionally unrelated proteins annotated as AtpI/UncI.

An interdependence of Fo subunits on each other for stability in the membrane after insertion has been shown. While Fo-b and Fo-c each insert into the membrane independent of the other two Fo subunits, Fo-a insertion is dependent upon co-insertion of Fo-b and Fo-c (122,203). Uncomplexed Fo-a does not persist in the membrane as it is degraded by FtsH (10). The Fo subunits may interact as a complex during insertion or alternatively Fo-a is incorporated into a bc complex already present in the membrane. In support of a dependence of Fo-a on other Fo subunits for membrane stability, Fo-a has been shown to interact with Fo-b and Fo-c during assembly (155,506). Although Fo-a stability in the membrane has been shown to be dependent on the other Fo subunits, overexpressed Fo-a can insert into the membrane (143,317,511). It is not know if the over-expressed Fo-a is incorporated into complexes with the other subunits of the ATP synthase.

Although subunits of the Fo domain interact during assembly, the targeting and insertion requirements are unique to each subunit. Fo-c does not require the Sec translocase for insertion into the membrane but rather depends solely on the membrane insertase YidC (499). Targeting of Fo-c to the membrane occurs independently of the SRP pathway (543). In contrast, Fo-a and Fo-b require the Sec translocase for insertion and are targeted to the membrane via the SRP pathway. Fo-a also has a strict requirement or YidC for efficient insertion (257,543). None of subunits of the Fo domain require the PMF For insertion.

The Fo part can assemble in vitro independently of the Fo part (451,452) and although some studies suggest assembly of Fo and F1 in vivo does not appear to occur independently (101,436) others have shown that ATPase activity due to partial or intact F1 complexes can accumulate in the cytoplasmic fraction of E. coli strains carrying mutant forms of ATP synthase (436). It has been proposed that interdependence of the Fo and F1 domains for assembly may ensure that an open proton pore does not exist as an intermediate (101). However, when the Fo subunits are expressed in the absence of the F1 domain, no increase in proton permeability of the membrane is observed (317) and only upon binding of the F1 domain does an active proton conducting pore form. The δ subunit as been specifically implicated in this activation (318).

The assembly of the F1Fo ATP synthase is illustrated in Figure 9. During assembly of ATP synthase α subunit must first be incorporated with β into the αβ3 hexagon before δ/α interaction can occur. Only after incorporation of α into the αβδ hexagon does the N-terminal region of α become available for binding δ (436).
Chapter 1

Assembly of the αβ hexagon also allows the assembly of the ε and γ into the F₁ domain. Interaction between the F₁ and Fₒ domains is mediated by F-β and F-ε with F-α not contributing significantly (259,506). The cytoplasmic domain of F-β cannot bind to soluble α (436), suggesting that the formation αβ hexagon is also a prerequisite of F-β binding. F-α is likely the last subunit to be added during ATP synthase assembly since in E. coli, truncated F-α (506) still allows the assembly of the enzyme complex and in Bacillus SP3, F-α-less ATP synthase has been isolated (350).

In yeast, Oxa1 has been implicated in the biogenesis of the F₁Fₐ ATP synthase (Complex V) as the complex is barely detectable in Δoxa1 cells and ATP synthase activity is strongly decreased (15). The actual role Oxa1 plays in Atp9 insertion appears to differ to that of YidC in F-ε insertion. Rather than playing an essential role in the insertion Atp9 into the membrane, Oxa1 interacts post-translationally with Atp9 to promote its assembly with Atp6 (225). This proposal is based on the observations that in the absence of Oxa1 Atp9 levels are actually less reduced than other F-ε subunits (225,272) and that inactivation of the inner membrane protease Yme1 in the Δoxa1 mutant restores oligomycin-sensitive ATP synthase activity and stabilizes the ATP synthase membrane subunits that would be otherwise degraded (272). Also, in contrast to the cytochrome c oxidase complex, activity and subunits of the F₁Fₐ ATP synthase are still detectable indicating that Oxa1 is not essential in the assembly of this complex (15). During assembly of the ATP synthase complex, newly synthesized Atp9 oligomerizes and forms an F₁-containing subcomplex before associating with the Atp6 (F-ε) protein. The inner membrane protein Atp10 acts as an Atp6-specific chaperone coordinating the assembly of Atp6 with the F₁-Atp9 subcomplex (204,484). Oxa1 is proposed to then interact with Atp9 coordinating the efficient and correct assembly of Atp9 oligomers so that they are competent for subsequent assembly with the F₁ subcomplex. A defect in the oligomerization of Atp9 with the F₁ subcomplex would therefore result in a failure of the Atp6 subunit to assemble with the other subunits of the F₁Fₐ ATP synthase resulting in a dramatic decrease in the amount of enzymatically active ATP synthase in the mitochondrion. The interaction of Atp9 with Oxa1 may ensure that the ATP synthase remains inactive i.e. that there is no proton leakage, until the complex is fully assembled. Oxa1 has been shown to copurify with the ATP synthase (225) and similarly in Bacillus subtilis, all subunits of the F₁Fₐ ATP synthase copurify with the YidC homologues SpoIIIJ and YqjG (402) suggesting a role for these bacterial proteins in the assembly of the ATP synthase.
FIGURE 9. Assembly of the F$_1$F$_0$ ATP synthase. The formation of the F$_1$α-F$_1$β heterohexamer renders the N-terminus of F$_1$α available to bind F$_1$δ in a stoichiometry of α$_3$β$_3$δ$_1$. F$_1$ε and F$_1$γ binding follows. F$_0$c forms a 10 subunit ring independently of the other subunits to which F$_0$b and F$_0$a assemble. F$_0$a is proposed to be the last subunit to assemble in the F$_1$F$_0$ ATP synthase and interaction between the F$_1$ and F$_0$ domains is mediated largely by F$_0$b and F$_0$c with little contribution by F$_0$a. The structure shown is a composite of four different X-ray crystallography and NMR spectroscopy structures from (519). The F$_1$-domain (PDB# 1C17) is connected to the F$_0$-domain (PDB# 1E79) by a stator consisting of F$_0$b (PDB# 2A7U) and F$_0$c (PDB# 112P).

14. CELLULAR LOCALIZATION
The traditional Fluid Mosaic Model of biological membranes is being replaced by the “Partitioned” or “Compartmentalized Fluid” Model (139, 263, 294). This new membrane model, supported by recent observations of supercomplexes and non-random protein distribution in the membranes of living cells, sees non-random distribution of membrane proteins as the norm rather than the exception. Relatively little is known about the supramolecular distribution of membrane proteins involved in energy transduction and two opposing models have been proposed to describe their distribution (reviewed in (274)). The Random Diffusion Model proposes that no large-scale organization of complexes exists while “solid state” models suggest that electron transport occurs within stable supercomplexes of respiratory proteins. In solid state models local non-equilibrium conditions exists resulting in heterogeneity. In support of the latter model, “respirasomes”, or respiratory supercomplexes, have been identified in a number of organisms, for
example *Paracoccus denitrificans* (422), *Corynebacterium glutamicum* (68) and the archaeon *Sulfolobus* sp. strain 7 (218) but to date only a dimeric *E. coli* cytochrome *o* oxidase has been reported (450). Structures of the identified supercomplexes have also been resolved by single-particle cryo-EM (129,421). In turn, it has been postulated that these respirasomes are located in specialized domains in the cell membrane forming “respirazones” (275). In *E. coli*, a GFP-fusion protein of CydB was expressed from its native locus and was observed to localize to mobile patches in the membrane (276). In the study it was estimated that the patches contained approximately 76 cytochrome *bd* complexes which would only account for 30% of the surface area of the membrane patches suggesting that other proteins are contained in these regions as well. In contrast, the largest biochemically characterized respiratory supercomplexes contain 4 to 6 electron transport complexes (57,423) indicating that the patches observed are not merely the supercomplexes reportedly previously. Preliminary results of the distribution of NDH-2, NDH-I, succinate dehydrogenase and the F$_1$F$_o$ ATP synthase shows that all these proteins are located in mobile patches in the membrane (276). It has yet to be determined whether these patches colocalize.

The respirazone hypothesis predicts that respiratory proteins are compartmentalized in the membrane and are segregated from other membrane functions. This is distinct from previous hypotheses that respiratory complexes are organized at the level of proteins and quinols. The physiological relevance of respirasomes and respirazones may be to improve the specificity of reactions of the quinol intermediates and prevent non-specific reduction of other membrane components. If the F$_1$F$_o$ ATP synthase is also located in the respirazones this could enhance the efficiency of the proton circuit as the proton pumps and sinks would be in close proximity preventing proton diffusion at the membrane surface which leads to PMF dissipation with increasing distance from the proton pump (92). There have been no studies as to what could hold the respirazones together. Cardiolipin, found exclusively in bacterial and mitochondrial inner membranes, has been shown to be involved in the organization of respirasomes in other organisms (363,549). In yeast, studies in mutants unable to produce cardiolipin have shown that this phospholipid is essential in the formation of Complex III and IV supercomplexes and does not merely increase their stability (549,550). The existence of lipid domains was suggested in 1982 by Karnovsky *et al.* (239) and since the membrane lipids of *E. coli* are heterogeneously distributed in the membrane, a network of protein-lipid and protein-protein interactions may be
involved in the formation of respirazones. Cardiolipin membrane domains have been visualized in *E. coli* by 10-N-nonyl acridine orange (NAO) staining and are observed at the septal region and at the cell pole (311). This suggests that cardiolipin is not involved in the formation of respirazones as it is with respiratory supercomplex formation but whether other phospholipids are involved in the segregation of the membrane into zones of specialized function has yet to be investigated.

15. CONCLUSIONS
Since coordination data on most respiratory complexes is lacking it is difficult to imagine how subunits located to the membrane via different translocases/insertases assemble together to form a functional complex of the correct stoichiometry and with the correct cofactor content. There appears to be a great amount of variance in the way in which chaperones, targeting pathways, and translocases coordinate membrane biogenesis and that each respiratory complex needs to be studied individually. This is illustrated by the difference in assembly structurally related succinate dehydrogenase and fumarate reductase as well as the absence of a universal maturation pathway for cytochromes (468). All insertion and translocation machinery available to *E. coli* is used for the insertion and assembly of respiratory complexes. In general metalloproteins to be exported to the periplasm achieve this via the Tat translocon once cofactor insertion, protein folding and, if relevant, subunit association has occurred. The presence of the highly conserved Tat signal peptide makes identification of such proteins possible. There are however metalloproteins that, through the interaction with another Tat signal peptide-containing subunit, can be translocated across the membrane even though they themselves do not contain a Tat-targeting sequence. For this “piggy-back” mechanism to occur, the Tat signal peptide must be “hidden” until cofactor insertion and subunit coupling has occurred, a role most likely performed by enzyme-specific chaperones. As illustrated by NarJ, chaperones that coordinate the cofactor insertion and coupling of soluble catalytic subunits, also coordinate the anchoring of these subcomplexes to the matured membrane anchor. Membrane anchors of respiratory complexes are often cytochromes which must undergo maturation in the membrane after insertion via the Sec translocase. There is no evidence to date that such chaperones can “sense” maturation of the membrane anchor subunits. Coordinated translation from a polycistronic messenger may contribute to correct timing of complex assembly although rates of translocation
via the Tat and Sec systems vary considerably. The rate of translocation via the Tat pathway in chloroplasts is in the order of nmol/min (323) while the *E. coli* Sec system translocates at pmol/min rates (472). Kinetic data on YidC-mediated insertion is missing.

Assembly of the larger respiratory complexes NDH-1 and the F$_{\text{f}}$F$_{\text{o}}$ ATP synthase are even more complex owing to the number of subunits and proton translocation capabilities of the proteins. An open pore in the membrane could not be tolerated during assembly. There does, however, appear to be assembly, albeit it partial in the case of the F$_{\text{o}}$-domain, into functional domains before the holoenzyme assembles. This may be a glimpse into the evolutionary development of these large respiratory complexes. Since bacterial respiratory complexes such as NDH-1, cytochrome $\text{o}$ oxidase and the F$_{\text{f}}$F$_{\text{o}}$ ATP synthase represent minimal forms of those present in the mitochondria, assembly processes present in bacteria may in turn provide insight into not only assembly of the core components of their eukaryotic counterparts but also the addition of subunits during the evolution of mitochondria. While our understanding of membrane biogenesis of respiratory enzymes is growing, evidence that respiratory complexes form supercomplexes and maybe even specialized domains in the membrane indicates another level of assembly to be explored.

16. SCOPE OF THIS THESIS

In this thesis, the role of the *Escherichia coli* YidC in membrane biogenesis is investigated with a focus on subunits of multisubunit respiratory complexes. **Chapter 2** describes a directed proteomics approach that was applied to investigate the effect of YidC depletion on components of the respiratory chains expressed under anaerobic conditions with either nitrate or fumarate as terminal electron acceptors. It was found that YidC depletion under these conditions caused a cessation in growth comparable to that seen in aerobically grown cells and that a number of respiratory complexes were negatively affected by YidC depletion. This was further investigated and is described in **Chapter 3**, where metabolic labeling with $^{15}\text{N}/^{14}\text{N}$ was employed to study the membrane proteome-wide changes upon YidC depletion under both aerobic and anaerobic respiratory conditions. The results obtained in Chapters 2 and 3 allowed the identification of putative YidC substrates and **Chapter 4** describes the characterization of the *in vitro* insertion requirements of subunit K of the NADH ubiquinone dehydrogenase. It was found that NuoK absolutely requires the Sec translocase and YidC for insertion and the
basis of the YidC-dependence was investigated focusing on essential membrane-embedded negative charges contained in NuoK. Chapter 5 describes the use of in vitro translation and membrane insertion systems to further characterize the role of YidC in the membrane biogenesis of the mechanosensitive channel of large conductance MscL. The study focuses on achieving fully active oligomeric MscL ion conducting channels in a cell-free system. YidC has been shown to copurify with the heterotrimeric SecDFYajC complex, and the effect of SecDFYajC depletion on the membrane proteome of *E. coli* was therefore investigated as described in Chapter 6. In Chapter 7 the findings are discussed in the context of our current knowledge of the Oxa1/Alb3/YidC family of proteins with particular focus on the diverse functioning of YidC in membrane biogenesis in *E. coli*.