Cytochrome c Biogenesis: Mechanisms for Covalent Modifications and Trafficking of Heme and for Heme-Iron Redox Control

Robert G. Kranz,1* Cynthia Richard-Fogal,1 John-Stephen Taylor,2 and Elaine R. Frawley1

Department of Biology1 and Department of Chemistry,2 Washington University, One Brookings Dr., St. Louis, Missouri 63130

INTRODUCTION

Heme is protoporphyrin IX with iron (Fig. 1). It is also sometimes called protoheme IX or heme b when it is associated as the prosthetic group of b-type cytochromes. A comprehensive review on the enzymes involved in heme biosynthesis and their regulation has been recently published (109). The present review covers covalent additions to the porphyrin of the heme molecule, with emphasis on heme c (cytochrome c assembly) and emerging concepts. In all cases discussed here, the natural modifications to heme occur after the ferrous iron assembly and the modifications to make various types, such as factors might be involved.

Heme has been used as a cofactor for over 3 billion years, and the modifications to make various types, such as c heme in a cytochrome c, have probably existed for almost as long. Consequently, these ancient molecules have stood the test of time. A cytochrome c has at least one thioether bond formed between a cysteine of the protein and a vinyl group of the heme. Almost always, two cysteines in a CXXCH “heme binding motif” of the protein are linked by two thioether bonds to the two vinyl groups of heme, where the histidine acts as an axial ligand to the heme iron (Fig. 1). In general, it is thought that such covalent linkages provide stability to the cytochrome, making them much less likely to lose their heme or denature (9). Stability might be particularly important to cytochromes which function outside of the inner membrane, where all cytochromes c are found. While the mitochondrial cytochrome c, containing a single heme group, has received the lion’s share of attention in the last century, it has become clear that in the prokaryotic world there are hundreds of cytochromes c with structures and functions unlike those of mitochondrial cytochrome c (128). Some of these may have more than 10 heme groups (each attached to a distinct CXXCH motif) in a single polypeptide. Moreover, they function in diverse and sometimes extraordinary electron transport chains that have evolved in an untold number of prokaryotic species. Many species have the classic aerobic respiratory chain from which the mitochondrion is derived, but countless cytochromes c are involved in photosynthesis or reduction of terminal electron acceptors such as nitrates, dimethyl sulfoxide, metals, and more (128). These oxidation/reduction reactions have a profound impact on the planet’s environment, including all of the earth’s nutrient cycles.

The typical purpose of these diverse electron transport systems is to couple electron movement across a membrane to the work of transport. The energy for this work is usually derived from the oxidation of electron donors such as glucose, pyruvate, or nitrites, and the cytochrome c proteins are components of the electron transport chain. Cytochromes c may use this energy to transport a range of molecules, including oxygen, protons, or electrons, and many species have cytochromes c that contain one or more active sites for reactions on the same polypeptide chain.

EVIDENCE FOR WHY THREE SYSTEMS EXIST: EVOLUTION AND ECOLOGY

FOR GENOME ANNOTATORS: WHAT DEFINES A SYSTEM I (ccm GENES), A SYSTEM II (ecs GENES), OR A SYSTEM III (CCHL)

REFERENCES

*Corresponding author. Mailing address: Department of Biology, Campus Box 1137, Washington University, One Brookings Dr., St. Louis, MO 63130. Phone: (314) 935-4278, Fax: (314) 935-4432. E-mail: kranz@biology.wustl.edu.
chains is the generation of a proton gradient across the inner membrane and thus energy conversion. This gradient is in turn used to make ATP, for motility, for transport, and for other processes common to all life. It is the respiratory diversity that has, in part, made the prokaryotic world so successful in inhabiting nearly every microenvironment on earth. For example, in *Shewanella* species, “nanowires” assembled with different multiheme cytochromes c shuttle electrons from organic substrates inside the bacterium to metal precipitates such as iron oxides or uranium outside (67, 143). In so doing, they often solubilize the metal. This molecular nanowire involves four different c-type cytochromes, three each with 10 c-type hemes (MtrA, MtrC, and OmcA) and one with 4 hemes (CymA) (143). Understanding the mechanisms by which organisms attach heme to each CXXCH motif in all cytochromes c is the major topic of this review. Of the three pathways to

<table>
<thead>
<tr>
<th>Heme structure</th>
<th>Genes for synthesis and comments</th>
</tr>
</thead>
</table>
| ![Heme b](image) | = heme  
= protoheme IX  
= protoporphyrin IX with iron  
Fe²⁺ inside the cell, with insertion by ferrochelatase as last step in biosynthesis (see ref 109 for genes and enzymes in heme synthesis). |
| ![Heme o](image) | Genes:  
*ctoB = cyoE = CQX10*  
Heme o synthase = farnesyl transferase  
CyoE has seven TMDs with most conserved residues on the neg side (inside) of the membrane. |
| ![Heme a](image) | Genes:  
*ctoA = CQX11*  
Heme a synthase (CtoA) has a b-type heme for oxidation reaction of substrate heme o (see C8 of heme a).  
CtoA has eight TMDs with most conserved residues on pos side (outside) of the membrane. |
| ![Heme c](image) | System I = ccmABCDEFGH  
System II = ccsB ccsA  
System III = CYC3 (CCHL) |

FIG. 1. Types of heme (b, o, a, and c), modifications and genes/enzymes required for modification, as discussed in the text. The Fischer numbering system was used.
assemble a cytochrome \(c\) (i.e., to covalently attach the heme), prokaryotes use the pathways called systems I and II. System I is comprised of eight proteins (CcmABCDEFGH) and system II of minimally two proteins (CesBA) that are dedicated strictly for cytochrome \(c\) biogenesis (Fig. 2). All are integral inner membrane proteins.

Here we bring in old and new ideas on mechanisms of cytochrome \(c\) assembly via analogies to other heme modification mechanisms. We propose future challenges in fully understanding the molecular mechanisms and evolutionary underpinnings behind the biosynthetic pathways. One of the goals of this review is to move from the traditional “protein blob” models to a membrane topological and structure/function level in describing the biosynthetic pathways. We believe that experimental topological data and recent biochemical studies of the systems facilitate this level of understanding.

HEMEN TYPES AND MODIFICATIONS

Experimentalists over the last century have characterized a handful of major types of heme molecules with various natural modifications (Fig. 1). These hemes were initially discovered based on their spectral characteristics and later characterized based on their unique heme chemistries that yield the distinct spectral absorptions. Over the last 2 decades, the biosynthetic mechanisms behind these heme modifications have been at the forefront (reviewed in, e.g., references 84, 115, and 154). Because there are some similarities among the chemical mechanisms and positions in the porphyrins for certain modifications, it is instructive to review the prominent types of hemes (Fig. 1).

Except for heme \(d\), which appears to have bireductive hydroxylations at carbons 5 and 6 of the porphyrin rings, other noteworthy modifications to heme involve additions to one or both vinyl groups at the porphyrin carbons designated 2 and/or 4 (Fig. 1). The modifications can range from a single change catalyzed by a single enzyme to a complicated pathway involving many protein accessory factors. In the former, hemes \(o\) and \(a\) both have a farnesyl addition to vinyl group 2, catalyzed by the protoheme IX farnesyltransferase enzyme, also called heme \(o\) synthase. Heme \(o\) synthase is encoded by the \(escherichia\) \(coli\) cyoE, the \(bacillus\) \(ctaB\), and the yeast \(COX10\) genes, all of which are related (133, 156). Additionally, in heme \(a\), a methyl group at carbon 8 is oxidized to an aldehyde. Heme \(a\) is the substrate for this modification, which is catalyzed by a heme-containing monoxygenase (or peroxidase) enzyme called heme \(a\) synthase, encoded by the \(bacillus\) \(ctaA\) and the related yeast \(COX11\) genes (31, 32, 104, 149–151). In vivo, the heme \(o\) and \(a\) synthases may form a complex for substrate channeling (30). Note that \(e.\) \(coli\) has only cyoE, and thus one of its two terminal oxidases is a cytochrome \(bo\) copper oxidase, which is related to the more publicized cytochrome \(aa3\) copper oxidase (35, 36, 119). The other \(e.\) \(coli\) oxidase is the unrelated noncopper cytochrome \(bd\) (69).

The enzymes for modifications in heme \(o\) and heme \(a\) (Fig. 1) show some parallels to certain features of cytochrome \(c\) synthesis that are worth mentioning. For example, in vitro addition of the farnesyl pyrophosphate substrate to heme by the cyoE-encoded heme \(o\) synthase (in vesicles) absolutely required heme-iron reduction by sodium dithionite (135). In that report, Saiki et al. stated, “The resulting polyenyl cation undergoes a nucleophilic attack by the vinyl group of ferrous protoheme IX, thus succeeding to a transfer of the farnesyl group to position 2 (beta) of the 2-vinyl group of ferrous protoheme IX with concomitant addition of a hydroxyl group to position 1. Since the electron density at the 2-position of the vinyl group is assumed to be higher in ferrous protoheme IX than in its ferric state, the former must be the real substrate for farnesylation.” Thus, as in cytochrome \(c\) synthesis, the heme iron must be in the reduced state so that the vinyl groups are chemically active for ligation (see below). Although the model described by Saiki et al. suggests that the resulting heme \(o\) product maintains iron in the reduced state after farnesylation, this was not tested because of the absolute requirement for dithionite in the in vitro reaction. CyoE was predicted and experimentally demonstrated to contain seven transmembrane domains (TMDs) (34–36). Most of the highly conserved residues are on the cytoplasmic side, where one would predict that reduced heme \(b\) enters for farnesylation. CyoE also contains a conserved binding site located in a cytoplasmic loop for poly-prenyl diphosphate. Saiki et al. showed the functional importance of key residues in CyoE at the cytoplasmic surface of the inner membrane, including residues that may be axial ligands to the heme iron, such as histidine and tyrosine (134).

The aforementioned heme \(a\) synthase (CtaA) is predicted to contain eight TMDs, interestingly, with five conserved histidines that are located on the periplasmic side of the membrane; these, in theory, are for coordinating two hemes that were shown to be present (76). One is a \(b\) heme that transfers electrons in the oxidative reaction, and the other is the heme \(o\) substrate (32, 149–151). Since the conserved histidines are at the outer surface of the membrane, it is suggested that the heme \(b\) substrate for CyoE (CtaB) and the heme \(o\) substrate for CtaA may be located on opposite sides of the inner membrane. As discussed below, certain analogies to CcmF in the system I cytochrome \(c\) assembly can be made for the CtaA heme \(a\) synthase.

Heme \(c\) synthesis requires involvement of the apocytochrome, that is, attachment of the CXXCH motif to the heme, so the terms “synthesis of (holo)cytochrome \(c\)” and “synthesis of heme \(c\)” are necessarily the same. Cytochrome \(c\) synthesis is by far the most complex heme modification, with three different systems having evolved to accomplish the same ligation (84, 115) (see below). The cysteine residues of the CXXCH motif must be reduced to form the thioether bonds to the alpha carbons of the heme vinyl groups. Some recent reviews on cytochrome \(c\) synthesis have included thorough discussions of the thiol reduction components (58, 157), and so the present review will only briefly describe these components. One aspect that makes cytochrome \(c\) synthesis even more phenomenal is that the polypeptide folds around the heme only after ligation, unlike the case for hemoglobins and other cytochromes, where protein folding occurs prior to heme binding. It is well documented that most apocytochromes \(c\) are degraded in vivo if the heme is not attached, yet apoglobins and apocytochromes \(b\) fold properly without heme and can be purified and reconstituted with heme in vitro (see reference 18 for a discussion).

Reactions of protein side chains to the heme vinyls have been studied in other contexts with various approaches, and the particular reactivities of the vinyl groups have been elegantly addressed, as have the ramifications of modifications
FIG. 2. Working models for cytochrome c biogenesis by systems I, II, and III. Models include trafficking and oxidation states of heme, as well as the subpathways for apocytochrome reduction (in red for system I and system II). Representative genera possessing each system are listed under the models.

**System I**

- **CCM**
- Signal sequence
- Apo-Cyt C
- Sec system
- CcmA
- CcmB
- CcmC
- CcmE
- Heme
- Heme
- CcmH
- Heme
- DsbD
- Outside (p-side)
- Inside (n-side)

**System II**

- **CCS**
- Signal sequence
- Apo-Cyt C
- Sec system
- CcxA
- CcsB
- CcsX
- Heme
- Heme
- CCHL
- Fungal mitochondria
- Invertebrate mitochondria
- Vertebrate mitochondria

**System III**

- **CCHL**
- Intermembrane space (p-side)
- Heme
- Fe²⁺
- Protoporphyrin
- Inside (n-side)

Representative genera possessing each system are listed under the models.
(78). For instance, Metcalfe et al. studied ascorbate peroxidase, which was mutated to form a spontaneous covalent bond to the 2-vinyl group (100). Interestingly, if the amino acid residue adjacent to the heme vinyl was changed to a cysteine, the covalent adduct required iron(II) (Fe$^{2+}$). If the residue was changed to a methionine, a covalent adduct required iron(III) (Fe$^{3+}$) for formation, and an oxidative H$_2$O$_2$-catalyzed reaction was proposed. Similarly, Barker and colleagues (19, 20) have analyzed a unique cytochrome residue adjacent to the heme vinyl was changed to a cysteine, dastase, which was mutated to form a spontaneous covalent bond 514 KRANZ ET AL.

The three systems for cytochrome c synthesis are summarized in Fig. 2, as are representative genera that possess each system. For systems I and II, there is a good understanding of how the CXXCH cysteines are reduced, using a general trans-membrane thioredoxin protein such as DsbD (101) (or the smaller, related CcdA [43, 139]) and dedicated thioredoxins such as CsxX (23) or CcmG (24). Figure 2 diagrams the thiol proteins (in red) which transfer reducing equivalents from inside the cell to outside. Since some aspects of assembly have been reviewed in the last 5 years (5, 9, 12, 58, 70, 157), here we focus on selected emerging concepts. Because all proteins in systems I and II are integral membrane proteins, descriptions will be presented in the context of the two-dimensional topologies of the integral membrane proteins in the pathways. Special emphasis will be placed on the completely conserved residues in this framework. For example, early studies of system I and II noted that three proteins possessed a highly conserved region called the “WWD domain,” due to the completely conserved tryptophans and an aspartate (WXWD) (25). It was proposed that this domain, in key proteins in the pathways (CcmC, CcmF, and CcsA), bound the heme that was presented to the apocytochrome c for ligation. In each protein, the WWD domain is flanked by conserved histidine residues originally predicted to ligand the heme iron and later shown to face the periplasm. The authors of a recent review on the evolution of the three members of this superfamily termed this the “heme-handling protein” family (94). The present review will spotlight the heme molecule as it moves through the systems, including the oxidation-reduction state, and it will include the very recent biochemical and spectroscopic evidence that these proteins do indeed handle heme, using the external histidines as axial ligands (59, 125). Our premise is that the WWD domain in each of the three families positions the heme molecule with the two vinyl groups surface exposed for orchestrated reactivity with the interacting partners. We next describe each of the three systems. For didactic purposes, we begin with the simplest, system III, and end with the most complicated pathway, system I. Some lessons learned about system II are also instructive for understanding system I.

**System III: CCHL**

In 1987, Dumont et al. reported the first identification of the gene encoding the enzyme that attaches heme to the two isoforms of yeast cytochrome c (49). This gene was able to complement the cytochrome c deficiency of the Saccharomyces cerevisiae CYC3 mutant, and it encoded a polypeptide with 293 amino acid residues. Designated the cytochrome c heme lyase (CCHL), it has since been shown to be related to the CCHLs of humans and some other eukaryotes (Fig. 2). CCHL has been recently reviewed (70), so here we provide only an overview.

A CCHL expressed with its cognate cytochrome c in the E. coli cytoplasm is all that is necessary to synthesize recombinant holocytochrome c, proving the enzyme’s cytochrome c synthetase activity (see, e.g., reference 118). [A note on enzyme nomenclature: currently it is common to use the term (holo) cytochrome c synthetase (or synthase) for the enzyme that directly catalyzes the heme attachment to apocytochrome c. As discussed above, the term CCHL was originally used in the yeast mitochondrial studies (see, e.g., reference 49). In this review we retain the use of CCHL for system III for historic and descriptive purposes. Clearly the CCHL is a cytochrome c synthetase.] In Fig. 2 we color code orange the proteins in each system which have been designated with the cytochrome c synthetase activity. Each cytochrome c synthetase must bind heme and the apocytochrome c substrates. In some organisms, such as yeast, there are two CCHLs, one for cytochrome c (CCHL) and a related one for cytochrome c$_1$ (CC1HL) (162). In other organisms, such as humans, a single CCHL recognizes and synthesizes both cytochromes c and c$_1$ (26). Thus, there appears to be more apoapocytochrome c substrate specificity than just the CXXCH motif for system III; systems I and II typically recognize only the CXXCH (see below). There must be features other than CXXCH within the cognate apocytochromes that the CCHL enzymes recognize. Defining the apocytochrome c substrate features and the CCHL recognition determinants remain important aims. Surprisingly, even though the CCHLs are soluble enzymes in the intermembrane space and have been functionally expressed in the E. coli cytoplasm, a structure has yet to emerge.

Although the CCHL gene in yeast was discovered 2 years earlier (48–50, 105–107) than the first system I ccm genes (85), there have been no reports of any other proteins necessary for system III, with the exception of Cyc2p (27). When the gene encoding Cyc2p is mutated, some cytochrome c$_2$ (27). Cyc2p is a flavoprotein, and it was proposed that it might reduce either the heme-iron or cysteines for ligation by the CCHL. It is also possible that dedicated disulfide-forming and thiol-reducing pathways are present in
mitochondria (17, 79, 82, 96, 144), but the exact mechanism(s) to maintain reduction of the apocytochrome c cytoenes for system III is unknown.

A heme binding site in CCHL has likewise been elusive. A particular sequence, CPX, was implicated as a heme binding motif and is required for CCHL function (145), but it has been noted that some CCHLs do not have this sequence (70). Also unknown is how the mitochondria employing system III traffic heme to the CCHL (Fig. 2) (99). Three possibilities come to mind: (i) a specific, yet-unidentified heme transporter such as described below for system II; (ii) simple diffusion through the inner membrane, with heme equilibrating between the membrane and intermembrane space, where CCHL binds it; and (iii) a broad substrate “flippase” like those used to “flip” phospholipids from the inside leaflet to the outside leaflet (33, 46, 77, 83), followed by binding to the CCHL. Recall that like phospholipids, heme is an amphipathic molecule (Fig. 1). Indirect support for the second or third mechanism comes from the heterologous synthesis of hemoglobin and cytochromes b in the E. coli periplasm (64). As long as the heme is maintained in the reduced state, the second or third mechanism could suffice for CCHL function and cytochrome c synthesis.

**System II: CcsB and CcsA Form a Heme Channel and Cytochrome c Synthetase**

The system II pathway emerged in the mid-1990s through genetic analyses of *Chlamydomonas* (47, 80, 159, 160) and selected bacteria such as *Bacillus* (92, 139, 140) and *Bordetella* (23, 55, 86) (Fig. 2). There were some inklings in 1992 that a system different from systems I or III existed when a comparison to the CcmC and CcmF proteins yielded a chloroplast open reading frame (ORF) which comprised a different WWD subfamily (25). This ORF was later called *ccsA* (cytochrome c synthesis) when it was deleted from the chloroplast genome of *Chlamydomonas*, resulting in a strain that could not make cytochrome *b* (160). Cytochrome *f* is analogous to cytochrome *c* of the *b* complex and contains the CXXCH motif. System II uses reducing equivalents from inside the cell to reduce DsbD (or a similar, smaller version called CcdA) (23, 139). A dedicated thioredoxin called CcsX is involved in reducing the cytochromes of the CXXCH motif directly (23). Note that exogenous thiol-reducing agents such as dithiothreitol can substitute for DsbD and CcsX activity; thus, they are not essential. The CcsA protein has the WWD domain and three conserved histidines, two predicted to be on the outside (H761 and H897) and one in TMD8 (H858). Concerning the topology of this protein, although hydropathy programs typically predict eight TMDs (94), two experimental analyses suggest six (62, 71). In the experimentally derived model, there are two hydrophobic patches that are placed in the bacterial periplasm rather than as TMDs, as shown in Fig. 3. The clear placement of these two conserved hydrophobic patches will await three-dimensional structural determination. Nevertheless, note that whether six or eight TMDs are present, no conserved residues change in location (topology), and so it is worthwhile to consider the roles of conserved residues in topological context.

The other absolutely essential protein for system II is called CcsB (80). In most organisms the CcsB and CcsA proteins are present as two separate polypeptides, usually encoded in an operon in bacteria. Experimental topological studies indicate that CcsB has four TMDs, with the third and fourth separated by a large periplasmic domain (200 to 500 residues) (Fig. 3) (23). Conserved residues in CcsB are very sparse, although the predicted secondary structures are maintained. The CcsB and CcsA proteins form a tight complex in *Bordetella pertussis*, and CcsB is completely degraded in the absence of CcsA (55). Likewise, in *ccsB* (*ccs1*) mutants of *Chlamydomonas*, the CcsA protein is not detectable (159). This may mean that critical interactions between TMDs of CcsB and CcsA exist. Early proposals suggested that the WWD domain bound heme, liganded by flanking histidines that would maintain the heme in the reduced state for presentation to the CXXCH motif (Fig. 3). The large periplasmic domain in CcsB would feature the apocytochrome c CXXCH recognition function, presenting the apocytochrome to the heme in CcsA.

In a few bacteria such as *Helicobacter*, *Bacteroides*, and *Wolinella*, the CcsB and CcsA proteins are naturally fused into one large ORF (called *ccsBA*). This ORF (from *Helicobacter*), when overexpressed in *E. coli* deleted of its *ccm* (system I) genes, was all that was necessary to synthesize a periplasmic cytochrome c (57). In this recombinant reconstitution, the CXXCH motif is reduced by the housekeeping reductants (e.g., DsbD and DsbC [E. R. Frawley and R. G. Kranz, unpublished data]). Thus, it is clear that CcsBA represents the system II cytochrome c synthetase. Topological studies of CcsBA give results similar to those for the separated CcsB and CcsA, as shown in Fig. 3, with the same caveat discussed above concerning the two hydrophobic patches in CcsA (59). Recently, Frawley and Kranz purified milligram quantities of the CcsBA protein with heme bound (59). Spectral analyses indicate that this heme is *b* type and is purified in the reduced state (Fe2+). Importantly, when the two flanking periplasmic histidines (H761 and H897) were individually mutated to alanines, the purified proteins still had heme, but it was in the oxidized state (Fe3+) and the spectral absorptions were significantly perturbed (59). We conclude that the function of each periplasmic histidine is to ligand the iron and maintain the heme in the reduced state for covalent thiether bond formation. This represents the “external heme binding domain,” which probably defines the active site of the CcsBA cytochrome c synthetase (Fig. 3).

The third conserved histidine (H858) in TMD8 near the bacterial cytoplasm was also mutated to an alanine. Notably, purified CcsBA (H858A) does not have heme in its external heme binding domain, yet the protein is stably synthesized (59). Another conserved histidine (H77) is located in the CcsB portion of CcsBA (and in all CcsB proteins). The CcsBA (H77A) protein contains fourfold less heme in its external heme binding domain. All four histidines in the recombinant CcsBA are required for synthesis of a cytochrome c (59), and they are also required in the *Chlamydomonas* CcsA (71) and CcsB (Ccs1) (47). Remarkably, the H77A and H858A TMD mutants are complemented for function by adding exogenous imidazole, but the histidine mutants that comprise the external heme binding domain are not (59). Complementation by imidazole has been demonstrated for recombinant myoglobin when the proximal histidine is mutated, wherein it was corrected for heme binding (21, 22). For this myoglobin “cavity
mutant, crystallographic results indicated that the cavity contains the exogenous imidazole, allowing the myoglobin to bind heme [22]. Our working model is that two heme binding sites are present in CcsBA, one at the channel entry (via H77 and H858) that facilitates delivery of reduced heme through a CcsBA channel, and a second site, the external heme binding domain, that is composed of the WWD domain and the two external histidine ligands that protect heme from oxidation. The results with CcsBA H77A and H858A also suggest that CcsB and CcsA are integrally associated, with a single heme binding site formed from the respective TMDs 3 and 8 (see Fig. 3).

Once the CXXCH motif is brought near the heme tethered within the WWD domain of CcsA, we predict that the spontaneous ligation to form the thioether bonds may induce the ligand switching to occur (i.e., the histidine of CXXCH for either the H761 or H897 in CcsA). Subsequently, we envision that this ligand switching and folding of the holocytochrome c polypeptide provides the energy to release the heme from the CcsA WWD domain. Note that one idea for system II is that it binds heme with a lower affinity than system I, so that the energy needed to release holocytochrome c may be less (see Evidence for Why Three Systems Exist: Evolution and Ecology below).

**System I: CcmABCDEFGH**

Genetic studies in the late 1980s and early 1990s started to unravel a complicated pathway for cytochrome c synthesis in a few model alphaproteobacteria such as *Rhodobacter* (24, 25, 85, 91), *Paracoccus* (111–114), and *Rhizobium* (120, 131). Although the players (now called the ccmABCDEFGH genes) were well defined genetically and were certainly dedicated specifically to cytochrome c synthesis, biochemical studies have only recently begun to elucidate the mechanisms involved. In short, cytochrome c synthesis by system I includes some fascinating chemistry and elegant molecular mechanisms for heme trafficking. It is instructive to describe system I as a two-step process, with oxidation and reduction of heme as it moves through the pathway (Fig. 2). One of the proteins in system I is a periplasmic heme chaperone called CcmE, discovered by Schulz et al. (142), and Reid et al. (122). Step 1 is the formation and release of oxidized (Fe³⁺) holocCmE (by CcmABCD), while step 2 ends with a cytochrome c synthetase.
complex (CcmF/H) that is proposed to use quinols to rereduce the iron of heme from holoCcmE, bring the CXXCH to this heme, and thus catalyze the ligation reaction.

**Step 1: heme delivery (to form oxidized holoCcmE) and the holoCcmE release subpathway, CcmABCD.**

Briefly, step 1 (Fig. 2) involves the flux of reduced heme (from ferrochelatase) to the periplasmic WWD domain of CcmC, concomitant with binding of CcmE to this heme/CcmC/CcmE ternary complex (Fig. 4 shows the proteins in step 1). The spontaneous covalent attachment of the 2-vinyl group of heme to the CcmE histidine (H130) then occurs (Fig. 4), which we refer to here as "adduct" formation (142). Adduct formation results in oxidized heme. A "release complex" comprised of the CcmABCD proteins uses ATP hydrolysis to release the oxidized (Fe$^{3+}$/H110) holoCcmE, freeing it to react with the CcmF/H complex (37, 56). CcmA is the ATP binding cassette (ABC) protein, while CcmB is a six-TMD protein (62) always encoded adjacent to ccmA (85, 120). Note that as diagrammed in Fig. 2, this step (formation of released holoCcmE) is likely very rapid in vivo. However, we have recently been able to trap, purify, and characterize intermediates in this step using various genetic strategies (125). These intermediates are reviewed here, and we discuss the chemistries underlying adduct formation and heme iron redox states.

(i) **Formation of the holoCcmCDE ‘preadduct’ intermediate complex.** It has been shown genetically that the six-TMD protein CcmC (Fig. 4) (62) is sufficient to bring heme to CcmE (56, 141). By removing the ccmA or ccmB genes, we have been able to trap heme in stable CcmCDE complexes (56, 125). Our recent results indicate that CcmC alone will not bind heme and that apoCcmE is necessary; heme is also necessary for CcmC to interact with apoCcmE, and so we propose that determinants on the heme molecule are recognized by both CcmC and CcmE (other than H130) are important for initial binding of heme. This is consistent with results of site-directed mutagenesis of CcmE,
where it was suggested that hydrophobic residues may be involved
in binding heme (54, 75). The intermediate holocCcmCDE complexes use two histidines in CcmC flanking the WWD domain as axial ligands to the heme (Fig. 4) (see below). Importantly, at least half of the heme in the preadduct holocCcmCDE complex is purified in the reduced state (125).

One of the important questions in the field has been how heme accesses each of its external sites (i.e., WWD domain), either through a protein channel (e.g., as shown above for CcsBA) or through the inner or outer leaflet of the membrane (see, e.g., references 62 and 155). For CcmC, there are no conserved histidines within TMDs, but other residues such as tyrosine are conserved (Fig. 4). Now that heme can be assayed in CcmC (as a complex) (125), it will be important to determine if a specific binding site in TMDs (and thus a channel from inside) is suggested. Because the covalent bond formation to CcmE H130 does not necessarily require Fe$^{3+}$ (see below), a nonspecific channel that does not protect the heme may be sufficient. Alternatively, without a need for protection from oxidation, heme could enter the CcmC WWD domain from the membrane directly.

(ii) Formation of the oxidized holocCcmCDE “postadduct” intermediate complex. Overexpression of the wild-type CcmCDE (no ccmAB) results in the postadduct intermediate complex, which has been purified and characterized (125). This holocCcmCDE complex contains all oxidized heme, with a striking UV/visible absorption spectrum when reduced with dithionite (125). A novel split alpha absorption is observed (at 553 and 559 nm), which was shown to be due to the unique chemical environment of the heme. Purification and spectral characterization of various CcmCDE mutants showed that the CcmE H130 covalent adduct and the CcmC H60 and H184 each contribute to the unique spectra. Spectra of complexes that contain alanine substitutions for either of the two external histidines in CcmC (H60 and H184) unequivocally show that these are the fifth and sixth axial ligands to the heme (125). Complexes with CcmC(H60A) and CcmC(H184A) have 50% less heme than the wild-type complex, and they exhibit major spectral perturbations. Interestingly, even with the CcmC(H60A) and CcmC(H184A) mutations, all of the heme is covalently linked to CcmE, consistent with the result that mutants with mutations in each were still partially functional (124). This also distinguishes the CcmE adduct formation from the CXXCH attachment noted earlier for analogous histidine substitutions in CcsBA. In CcsBA, the flanking histidine mutations completely abolish function, and these mutant CcsBA proteins were unable to protect the heme from oxidation (59). In this case, the wild type and CcmC histidine mutants result in an oxidized holocCcmCDE complex.

The covalent bond formed between the CcmE H130 and heme has received much attention. The unique chemistry and structure of this bond to the 2-vinyl β carbon (93) have been discussed (61, 93, 147, 148, 158). We discovered that all CcmCDE postadduct complexes and released holocCmE (see below) contain Fe$^{3+}$ heme, suggesting that in addition to covalent bond formation in CcmE, there must also be an oxidation step leading to an oxidized holoCcmE. Ferguson and colleagues have discussed the idea that the CcmE H130 imidazole would require a radical intermediate to form the covalent bond at the β carbon of the heme vinyl group (3, 93). However, the final oxidation state of the heme iron has drawn little attention. Figure 5A illustrates possible mechanisms for the adduct formation to CcmE H130. The first uses the H130 imidazole cation radical, but formation of this radical would itself require an oxidant (marked “Q” in the Fig. 5A, but its identity is unknown). The second mechanism is a nucleophilic reaction that is favored by an oxidized heme iron. The fact that heme in the preadduct CcmCDE complex is approximately 50% reduced (125) could argue for either mechanism. The in vitro synthesis of the holocCmE+ adduct appears to require the reductant dithionite (15, 42, 148), leading to the suggestion of a radical mechanism, although it is unclear whether a cation or an anion radical is proposed. The mechanism by which this radical might be formed in vitro or in vivo is still a mystery, but note that with the cation radical, the end product has oxidized (Fe$^{3+}$) heme. It seems unlikely the imidazole could form a cation radical in the presence of the dithionite reductant (since the oxidation potential of heme is around $-0.2$ V and that of imidazole is around $+0.5$ to 0.8 V at neutral pH). Imidazole reduction by dithionite to form an anion radical also seems unlikely. The alternative is to form the CcmE H130 adduct through a nucleophilic Michael addition, which would prefer oxidized (Fe$^{3+}$) heme (Fig. 5A). Likewise, the oxidant to form Fe$^{3+}$ heme, whether a quinone or semiquinone or other, is unknown. It is feasible that both mechanisms operate in vivo and that the partial function of the CcmC histidine mutants noted above reflects this plasticity. Clearly, the more likely chemical mechanisms for adduct formation to the β carbon predict a holoCmE with oxidized heme, and the data on intermediate holocCmCDE complexes and released holocCmE (125) are in agreement with these mechanisms.

Mutation of tyrosine 134 in CcmE(Y134A) was shown previously to have a dramatic effect on the spectra of purified holocCmE+ (no TMD), and thus tyrosine 134 is an axial ligand in released CcmE (158). In the intermediate CcmCDE complex with heme, a CcmE(Y134A) mutant showed a spectrum no different from that of the wild-type CcmCDE complex. This indicates that axial ligands to the heme iron have not yet switched over to the CcmE tyrosine 134 in the intermediate complexes, and thus all heme is still liganded to the CcmC H60 and H184.

(iii) Release of holocCmE via the CmABC complex. We conclude that the stable intermediate holocCmCDE complexes described above require the release function provided by CcmAB and that in the release process axial ligands must switch from CcmC (H60 and H184) to CcmE (Y134) (125). When the CcmA and CcmB proteins are coexpressed, the released holocCmE free from CcmC (56). We envision that major conformational movements are needed to accomplish this switching and holocCmE release, thus explaining the requirement for an ABC transporter complex (125). Note that a CmABC complex has been suggested for Rhodobacter capsulatus (62, 63, 65) and E. coli (56), but others in the field offer alternative views proposing that a CmAB transporter operates independently of CcmC by exporting an unknown substrate such as a reductant (37, 141, 146).

We believe that all data are consistent with the ABC transporter release function for CmABC (Fig. 6). Conceptually, the release is similar to that for the Lol ABC transporter (LolCDE), designated the lol system, that has been character-
We note that both the LolCDE proteins and the CcmABC proteins are not typical ABC transporters in that they do not translocate a substrate but instead are involved in recognition and release of periplasmic molecules. The release function of each has been shown to require the ABC component (LolD or CcmA) for a proposed conformation change and “release” of the target molecule, lipoproteins for LolCDE and holoCcmE for CcmABC (56). The lol system is dedicated toward recognizing lipoproteins that will be sent from the cytoplasmic membrane to the outer membrane of gram-negative bacteria. Lipoproteins have a fatty acid attached to an N-terminal cysteine residue. Some lipoproteins function tethered to the cytoplasmic membrane, but most require tethering to the inner surface of the outer membrane. LolCDE in *E. coli* recognizes lipoproteins that do not have an aspartate next to the cysteinyl attachment (97). The aspartate is an exclusion feature, and therefore lipoproteins with the aspartate remain as cytoplasmic membrane-anchored lipoproteins. Outer membrane-targeted lipoproteins are tightly bound to LolCDE in the absence of ATP.
hydrolysis (by LolD) or in the absence of the periplasmic LolA chaperone. Upon binding of LolA, ATP hydrolysis releases the lipoprotein to LolA, which transfers it to the outer membrane LolB. LolB and LolA are related, and this energy-independent exchange has been discussed previously (110). Thus, the holoCcmE release is analogous to the lipoprotein-LolA release. LolA is the lipoprotein chaperone to the outer membrane, and CcmE is the heme chaperone to what we refer here as “step 2,” the synthetase component of system I (Fig. 6).

It was previously determined that the small polypeptide CcmD is part of a CcmABCD complex in _R. capsulatus_ (63) and in _E. coli_ (127). Recently we have shown that CcmD is essential for the holoCcmE release but not for the heme attachment to CcmE or binding of CcmE to CcmC (127). As we have described for the _R. capsulatus_ CcmD, the _E. coli_ CcmD has a single TMD with an N terminus in the periplasm (Fig. 4). The short N terminus contains a completely conserved tyrosine of unknown function, although it was proven not to be an axial ligand in the intermediate CcmCDE complexes (125).

In summary, three recent experimental results lead us to suggest that the released holoCcmE is in the oxidized Fe$^{3+}$ state rather than Fe$^{2+}$: (i) all heme in the purified, trapped holoCcmE complex and holoCcmE is Fe$^{3+}$ (125); (ii) all heme in the purified WWD protein CcsBA from system II is in the reduced (Fe$^{2+}$) state, as one would predict, and thus even purification under ambient air yields a reduced WWD holo-protein (59); and (iii) a highly conserved mechanism to reduce the heme iron of holoCcmE appears to be present in CcmF [see “Step 2: heme attachment to apocytochrome c (from holoCcmE), and CcmFH as a holoCcmE reductase and cytochrome c synthetase” below] (125). (Daltrop et al. have observed heme transfer from holoCcmE* [truncated] to an apocytochrome c in vitro [albeit without CcmF/H present and after overnight incubation] (42). The heme for this transfer from holoCcmE* required reduction by dithionite. Although the authors concluded that one function of holoCcmE may be to retain the heme in the reduced state, unlike our proposal here, their data are consistent with a need for iron reduction in holoCcmE.)

**Step 2: heme attachment to apocytochrome c (from holoCcmE), and CcmFH as a holoCcmE reductase and cytochrome c synthetase.** Released holoCcmE is free to interact with the CcmF/H complex, but a key question is the mechanism by which the heme is reduced for covalent bond formation to the CXXCH motif. We have recently purified CcmF in milligram quantities from _E. coli_ (125). Similar to the results of Sanders et al. with _R. capsulatus_ CcmF (138), the _E. coli_ CcmF and CcmH proteins form a complex. Surprisingly, the preparations were tinted red, which is suggestive of heme (125). Initially, we thought that this might originate from holoCcmE, since it has previously been shown using coimmunoprecipitation experiments that CcmE and CcmF interact (2, 123). However, spectral analysis, heme assays, and pyridine hemochromagens indicated that this is a b-type heme and that (holo)CcmE was not copurifying (125). Even in the absence of the _ccmE_ or _ccmH_ gene, CcmF has a b heme in a 1:1 stoichiometry. Thus, the data indicate that the CcmF/H complex is itself a cytochrome b that has the cytochrome c synthetase function. This result is conceptually reminiscent of the heme a synthase encoded by _ctaA_ described above. Recall that CtaA has a heme b which is involved in electron transfer and the heme o substrate for modification to heme a (76). Based on experimentally determined topological models for CcmF (62, 129) and new _phoA_
fusion results (125), a unified topological diagram for CcmF is shown in Fig. 7. The completely conserved residues (in red in Fig. 7) are used to discuss the following three functions for the CcmF/H complex.

First, at least one histidine (CcmF H261) acts as an axial ligand to the heme to yield the cytochrome b function. A CcmF (H261A) mutant protein has one-quarter the amount of heme in the wild type and a perturbed absorption spectra (125). The other heme axial ligand could be H491, but this has not yet been tested. The heme of CcmF is reduced in vitro by quinols, suggesting that its cytochrome b functions as a quinol:heme oxidoreductase. We note a partially conserved motif, (“Q” in Fig. 7) with the amino acid sequence NPF at the outer surface of the second periplasmic loop, which is analogous to the quinone binding motif of RegB (GGXXNPF) in a comparable topological position (152). We suggest that the CcmF protein accepts electrons from quinols to reduce the heme, which subsequently reduces the heme from holoCcmE (125). Note also that the heme is positioned topologically directly below the WWD domain, the predicted location for the heme of holoCcmE (Fig. 7). Although we cannot rule out that the cytochrome b is directly involved in the thioether bond formation, the simplest interpretation is that it acts as a quinol:heme oxidoreductase in vivo.

A second predicted function of CcmF is binding to holoCcmE as part of its synthetase activity, using the holoCcmE heme for attachment to apocytochrome c (57, 123, 125). The CcmF (H173A) protein has wild-type cytochrome b function (125) yet is completely nonfunctional (123, 125). In analogy to CcmC and CcsA, as discussed above, the H173 and H303 of CcmF are periplasmically located, adjacent to the WWD domain, and are likely ligands to the heme from holoCcmE, which is used for the synthetase reaction (Fig. 7). How does axial ligand switching of this heme (from holoCcmE to CcmF to CXXCH) occur? Possibly, the energy from cytochrome electron transfer is used for this, and/or the covalent attachment to CXXCH induces this switching. Uchida et al. used Raman spectroscopy to indicate that one axial ligand in holoCcmE* (truncated with no TMD) is a histidine and the other is the tyrosine 134 discussed above (Fig. 4) (158), but it is not necessary to envision two heme ligands from CcmE immediately switched to two from CcmF and then to the two from cytochrome c. Garcia-Rubio et al. used various spectral approaches on holoCcmE* indicating that tyrosine is one axial ligand, but an unknown sixth ligand may exhibit conformational flexibility (61). In any event, it is likely that the WWD domain binds holoCcmE heme, precisely positioning the vinyl side chains (one with the CcmE adduct) to accept the CXXCH motif of the apocytochrome c (see below). The synthetase reaction is likely spontaneous, as long as the iron of heme is reduced (see Fig. 5B for chemical reaction of one thiol of CXXCH to the α carbon of the heme vinyl group). At this site, the CcmE H130 imidazole side chain must be ejected from the carbon of the 2-vinyl group. As shown in Fig. 5A (large blue arrow), iron reduc-
tion of the heme in holoCcmE would also favor ejection of the imidazole by the reverse Michael reaction.

The third function of the CcmF/H complex is recognition of the CXXCH motif for the synthetase activity. The apocytochrome c CXXCH motif may be recognized by CcmH, potentially chaperoning the motif to the WWD active site containing heme (from holoCcmE). Although the large periplasmic N- and C-terminal loops of CcmF (Fig. 7) have no designated functions as of now, possibly these are involved in interaction with CcmH. Both the R. capsulatus (138) and E. coli CcmF proteins are easily dissociated from CcmH, in the latter case simply by using high-salt conditions (125). These results suggest that interaction between TMDs of CcmF and -H are minimal and that electrostatic interactions via soluble domains may dominate, hence leading to the suggestion that the large periplasmic domains of CcmF are involved. CcmH has an RCXXC motif (Fig. 7), which was previously shown to reduce the CXXCH of apocytochrome c (see, e.g., reference 103). The structure of the first CcmH periplasmic domain (from Pseudomonas) containing the RCXXC motif has been shown to be a three-helix bundle, unlike the case for other thioredoxin-like proteins (45). The recently described analogous E. coli CcmH domain suggests that it is a dimer (1). Note that in most organisms (e.g., R. capsulatus) capsulatus is split into two ORFs (Fig. 7). The first was previously called ccl2 and has a single TMD and the RCXXC periplasmic domain (25, 103). The second ORF is called cycH (or ccmI), which had been topologically dissected and studied in alphaproteobacteria such as R. capsulatus and is potentially involved in the chaperone function (of the apocytochrome c), using a tetratricopeptide repeat (38, 44, 91, 136, 137). Employing two-hybrid analysis in yeast, Meyer et al. have shown that the Arabidopsis CcmH interacts directly with apocytochrome c (102). Because only the N-terminal periplasmic domain (Ccd2) appears to be essential for cytochrome c synthesis under certain conditions (see, e.g., references 44 and 132), it is important to key in on this domain (Fig. 7). It clearly has both chaperonin and thiol reduction functions. After attachment to the CXXCH motif, the heme must be released from CcmF/H. Potentially, the reactions of the two vinyl groups with CXXCH (including ejection of the histidine-vinyl from CcmE) and/or the subsequent folding of the apocytochrome c around the heme provides the energy for release from the CcmF/H active site.

When we first presented the sequence of CcmF (called Cc1) in 1992 (25) and subsequently performed a topological and residue consensus profile in 1998 (62), we wondered why CcmF was so large and why there were so many completely conserved histidines. We believe that multiple functions described here provide the rationale for these conservations and observations. Interestingly, in plant mitochondrial genomes, CcmF is separated as two (72) and sometimes three individual ORFs, called CcmF N1, CcmF N2, and CcmF C, which form a complex with CcmH (70, 121). CcmF has genomic organization of the mitochondrial endosymbiotic naturally separated multiple functions of the CcmF protein, albeit while still forming a complex? Finally, we must note that the detailed mechanism of formation of the two thioether bonds at the CcmF/H active site remains a very significant open question in the field. With the likely addition now of a mechanism to reduce holoCcmE at the active site prior to ligation, the molecular mechanisms and timing of reactions become even more intricate. Below we address why such a complicated system has been maintained for billions of years through evolution.

System IV and Others

System IV. Kuras et al. identified four integral membrane proteins (called Ccb1 to -4) that are necessary to attach heme covalently to a single cysteine in cytochrome b or of the cytochrome b complex (hence the designation ccb genes) (87, 88). The ccb genes are present only in oxygenic phototrophs (95), where the cytochrome b has been known for some time to have this spectrally unique heme. The cysteine is located on the opposite side of the membrane from the cytochrome f (c type), and so this putative Ccb complex functions on the opposite side (n side) from that of the three systems for cytochrome c synthesis (p side) just described. The authors suggested designating Ccb “system IV,” which seems appropriate. Like systems I to III, it is widely distributed, and clearly the mechanism of attachment and the functions of the Ccb proteins are of tremendous interest and may provide parallels with and clues to other systems down the road.

Others. The apocytochrome c recognition motif for system I has been much studied and appears to be only the CXXCH motif (6–8, 13, 14, 29), with the caveat that CXXXCH may also be used (130). Some variants of system I (e.g., NrfEFG [51], which is related to CcmFGH) and system II (CcBA) recognize slightly altered apocytochrome c signature motifs such as CXXCK (51, 74, 117) or CX15CH (73). These variants are likely to play major roles in the future in understanding of the apocytochrome c recognition determinants and the features of the substrate and synthetases that are important for such.

Some variations of system I exist in archaea and a few bacterial outliers (e.g., Desulfovibrio) in that the CcmE protein has a cysteine instead of the histidine (H130) for covalent bond formation (11). Note that a cysteine (thiol) can carry out reactions similar to those of the imidazole (histidine), forming a covalent bond to the β vinyl group of heme (Fig. 5A); likewise, this thiol adduct would be ejected preferentially by reduction of the heme iron (Fig. 5A). Interestingly, the CcmH protein is missing in the organisms that have a cysteine in CcmE. The significance of this should become clearer as we understand the chemical mechanisms underlying holoCcmE adduct formation and the CcmF synthetase reactions (on the holoCcmE and CXXCH motif of the apocytochromes). One speculation might be that the ejected thiol (of CcmE) reduces the CXXCH (rather than the CcmH RCXXC), and thus CcmH is no longer required for reduction, possibly with CcmF taking over the CXXCH recognition function.

Certain rare protozoans (Euglenozoa such as trypanosomes and Leishmania) have a cytochrome c in their mitochondria with a single cysteine in a CH motif (rather than the classic CXXCH) (10, 12). Surprisingly, they do not appear to have any proteins from system I, II, or III. The structure of their cytochrome c is similar to that of other mitochondrial cytochromes c, so Fulop et al. have speculated that a new, yet-unknown system is still to be discovered for these protozoa (60). It also remains possible that a diverged CHL that cannot be detected by BLAST searches is present yet recognizes more than just the CH in the apocytochromes. Recall that there clearly
are determinants on apocytochromes $c$ and $c_1$ (other than the CXXCH) recognized by the CCHLs. Possibly both the recognition determinants and CCHLs have coevolved, drifting so that they are not detectable by homology searches.

**EVIDENCE FOR WHY THREE SYSTEMS EXIST: EVOLUTION AND ECOLOGY**

Why are there cytochromes $c$? Studies in the last decade have suggested that the covalent linkage provides added stability to the cytochrome $c$, and so they are much less likely to lose the heme than the globins or other cytochromes (see, e.g., reference 16; reviewed in reference 3). The idea that a wider range of redox potentials result from attachment (28) has been referenced in reference 16; reviewed in reference 3). The idea that a wider range of redox potentials result from attachment (28) has been referenced in reference 16; reviewed in reference 3). The idea that a wider range of redox potentials result from attachment (28) has been referenced in reference 16; reviewed in reference 3).

Remember also that cytochromes $c$ in prokaryotes function outside the protected cytoplasm. Using cytochromes $c$ may be one way to “hard-wire” a chain or conduit for electron transfer that will function under diverse and sometimes hazardous environmental conditions.

Why are there three predominant systems for assembly rather than one? Let us start with system III, the simplest, which uses the CCHL enzyme. Clearly the CCHL evolved after mitochondrial endosymbiosis (which occurred with an alphaproteobacterium possessing the system I pathway) (68, 89, 90). Recall that all alphaproteobacteria possess system I, as do the mitochondria of plants and some primitive protozoa (Fig. 2). One case for the need for systems I and II in bacteria and plants is that they must ultimately end with a reduced heme iron for ligation to occur, even under environmental conditions that may be oxidizing. What would make the more complicated system I obsolete in certain mitochondria? For some eu-karyotes, it is likely that the internal milieu requiring synthesis by system I changed. We speculate that the heme was largely protected against oxidation, and additionally that reduced iron, and therefore reduced heme, was not limiting for growth (see below for a discussion of the benefits of system I). For example, one could envision that the levels of iron needed for Fe/S proteins, hemoglobins, and/or other cytochromes significantly surpassed that needed for cytochromes $c$ and $c_1$. A simpler mechanism that did not involve the synthesis of at least eight integral membrane proteins and the use of ATP (as with system I) became more economical. Since only two cytochromes $c$ were used (cytochromes $c$ and $c_1$), a specialized enzyme that recognizes these two became possible. Circumstantial proof for these ideas is the retention of system I in plant and some protozoal mitochondria, which likely face more oxidizing and varying environments. Possibly, some protozoa in the future with both systems I and III, in the process of losing system I and/or using each under different conditions, will be discovered.

At present, because of the significant amount of lateral transfer that has occurred with systems I and II, it is not possible to say which came first (Fig. 2) (66). Because of its simplicity, one would guess that system II evolved initially. Its presence in cyanobacteria and plant chloroplasts indicates that it was also part of the cyanobacterial endosymbiotic events leading to the algal chloroplast. Its retention might well be due to the fact that the heme is well protected from oxidation (59). Still, some accessory players in chloroplasts, possibly for maintaining reduced thios of the CXXCH in cytochrome $f$, may have evolved subsequent to the endosymbiosis. These have been discussed previously (116).

What possible advantages might system I confer that are not features of system II? Why is system I so complicated, as it clearly takes more energy to produce the system I proteins than system II and we know that ATP is required for system I function? We have evidence from the comparison of these two recombinant systems in *E. coli* that system I can use heme at potentially five-times-lower concentrations than system II (126). Iron is the limiting factor in heme biosynthesis (109). Thus, in certain iron-limiting environments an organism with system I can make a cytochrome $c$ while those with system II cannot. Moreover, in experiments where holoCcME was synthesized and then heme synthesis shut off completely (either with heme synthesis inhibitors or by genetically knocking out heme biosynthesis), *E. coli* with system I but not with system II was still able to synthesize cytochromes $c$ de novo (57). This means that holoCcME can function as a heme reservoir when no iron is available. (These experiments also prove that the heme attached to the cytochrome $c$ comes from holoCcME.) Together, these advantages might provide for a fitter organism in particular habitats. Interestingly, a very few species of bacteria have both systems I and II, such as *Bordetella parapertussis*. Possibly these are in the process of lateral transfer, awaiting selection against the system that provides less advantage. (For example, *B. pertussis* has only system II.) Alternatively, *B. parapertussis* may use system I under one growth condition and system II under a different condition. Studying the use of each system in these organisms and their ecology may further address the advantages and disadvantages of each in nature. We have attempted to evaluate the types of organisms that possess each system and to analyze the habitats in which these are typically found (B. San Francisco and R. G. Kranz, unpublished data). No clear patterns have emerged in this analysis, in part because the microorganisms that microorganisms use and their ecological distributions are poorly understood.

**FOR GENOME ANNOTATORS: WHAT DEFINES A SYSTEM I (ccm GENES), A SYSTEM II (ccs GENES), OR A SYSTEM III (CCHL)**

Given the fact that some proteins from both systems I and II (CcSA, CcmC, and CcmF) possess the highly conserved WWD domain, it can sometimes be confusing and time-consuming to properly annotate the sequences of new genomes for cytochrome $c$ synthesis proteins. What signature genes and sequences can identify a system when the picture is murky? Here we provide some brief guidance for proper annotation and how to quickly establish whether the organism has system I (ccm), system II (ccs), both systems I and II (rarely), or system III. Table I presents some signature motifs found in each of the subfamilies of the WWD proteins (heme-handling proteins) and other proteins. The signature genes for system I that should emerge from a BLAST search are *ccmC*, *ccmF*, and *ccmE*. In outlying organisms with system I, the other *ccm* genes
<table>
<thead>
<tr>
<th>System</th>
<th>Gene product</th>
<th>Conserved motif(s)</th>
<th>Additional key residues</th>
<th>Location of N terminus</th>
<th>No. of TMDs</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>CcmC</td>
<td>-WGX(F/W/Y)WXW(D/E)XR-</td>
<td>Two conserved His residues flanking the WWD domain</td>
<td>Inside</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CcmD</td>
<td>-(G/A)XY-</td>
<td>Inside 1</td>
<td></td>
<td>1</td>
<td>CcmD is always downstream of CcmC if it is present in bacteria; CcmD often shows as a small hypothetical protein of 40-80 amino acids; most organisms have the conserved -GXY- motif, with the exception of some alphaproteobacteria which have -AXY-</td>
</tr>
<tr>
<td></td>
<td>CcmE</td>
<td>-(C/H)XXX-</td>
<td>Inside 1</td>
<td></td>
<td>1</td>
<td>Most organisms have a His in the conserved motif; some archaea and deinococci have a Cys instead; Has up to 15 predicted TMDs (see Fig. 3B legend)</td>
</tr>
<tr>
<td></td>
<td>CcmF</td>
<td>-W(S/A)YXXLGWG[F/W/Y]WXWXDPVNAS<a href="L/M/I">F/L</a>WL- -LRCXXC- (thioredox domain), TPR repeat domains</td>
<td>Four conserved His residues</td>
<td>Outside</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CcmH/CcmI</td>
<td>-LRCXXC- (thioredox domain), TPR repeat domains</td>
<td>Outside 2</td>
<td></td>
<td>2</td>
<td>In most organisms, the E. coli CcmH is divided into two proteins, CcmH (Ccl2) and CcmI (CycH); CcmH is amino acids 1-127 of the E. coli protein and has one TMD and contains the -LRCXXC- motif; CcmI is a TPR repeat protein with one TMD; inability to identify CcmH does not automatically exclude an organism from being classified as system I</td>
</tr>
<tr>
<td>II</td>
<td>CcsA</td>
<td>-WAXX(A/S)WGX(F/Y)WXWDXK EXX-</td>
<td>Three conserved His residues, two flanking the WWD domain and one in a TMD on the n side</td>
<td>Outside</td>
<td>6</td>
<td>CcsA proteins may have 8 predicted TMDs (see text); they are also preceded or followed by CcsB in bacterial genomes</td>
</tr>
<tr>
<td></td>
<td>CcsB</td>
<td>Large soluble domain between TMDs 3 and 4</td>
<td>-VNXP- in soluble domain, conserved His in third TMD</td>
<td>Inside</td>
<td>4</td>
<td>CcsB proteins may have 5-6 predicted TMDs, due to hydrophobic regions in the extracellular domain; CcsB proteins have little sequence conservation, though the protein secondary structure is conserved, and they may proceed or follow CcsA in the genome</td>
</tr>
<tr>
<td></td>
<td>CcsBA</td>
<td>-WAXX(A/S)WGX(F/Y)WXWDXK EXX-</td>
<td>-VNXXP- in CcsB soluble domain, four conserved His residues (His-77, -716, -858, and -897; see Fig. 5)</td>
<td>Inside</td>
<td>10</td>
<td>CcsBA proteins may have up to 11-14 predicted TMDs (see text); CcsA-like proteins with 11-14 predicted TMDs that are over 2.5 kb in length may be candidate CcsBA fusions, particularly if a superfamily search indicates the presence of a ResB superfamily domain; a homology search against a known CcsBA fusion sequence (i.e., Helicobacter) can provide additional confirmation</td>
</tr>
<tr>
<td>III</td>
<td>CCHL</td>
<td>-CPX- (in N-terminal part), -PFDR HDW- (in C-terminal part)</td>
<td>See Pfam 01265</td>
<td>Mitochondrial intermembrane space</td>
<td>0</td>
<td>CCHL proteins are soluble proteins that lack a typical mitochondrial targeting sequence; The N termini are poorly conserved but contain several -CPX- heme-responsive motifs (with two known CCHL exceptions); proteins that group to Pfam 01265 and contain -CPX- motifs are likely CCHL</td>
</tr>
</tbody>
</table>
are most likely present, with the exception of $c$cmH, as noted above. For system II, the signature gene is $ccsA$ with the indicated WWD motif. As shown for the few conserved residues in CcsB (Fig. 3, in red), the $c$csB ORF can be troublesome to detect by homology searches in outliers, although it is typically in an operon with $ccsA$. For system III in mitochondria, all CCHLs are related, but specifically conserved residues in outliers also might be difficult to detect by BLAST searches (as discussed above for Euglenozoa). Finally, a few organisms can have multiple copies of a system. For example, Salmonella enterica serovar Typhimurium has two complete sets of $ccm$ genes (98). Remember that the variations that recognize rarer apocytochrome substrates may be additionally present (e.g., in Wolinella two more CcsBA proteins, one for a CXXCK motif, and in E. coli the $m$tf $g$ genes, which encode CcmFGH paralogs that recognize CXXCK).

**FUTURE STUDIES**

Throughout this review we have noted some particularly important studies that are needed in the field. Recall that the structures for only a few proteins in systems I and II have been determined, and then only soluble, truncated versions that are typically nonfunctional ($Cc$me*, $Cc$mg*, $CcsX*$ [ResA], and $Cc$mh* [periplasmic domains] [see, e.g., references 52 and 53] and others cited above). The crystal structure of a WWD domain (heme-handling) protein ($Cc$mC, $Cc$F, or $Cc$sA), with and without heme, remains a critical need. Although some in the field have proposed that the WWD domain interacts with the CcmE polypeptide component (70, 124), the presence of the WWD domain in CcsA (system II), where no CcmE operates, suggests otherwise. We suggest that the fundamental purpose of the WWD domain is as an external scaffold for heme, whereby the vinyl groups are surface exposed for interaction with various proteins, yielding a heme environment (including the flanking histidines as axial ligands to the iron) that is ideal for the necessary vinyl reactivity. The three-dimensional structures of the heme $o$ and $a$ synthases are important in and of themselves but also for comparison with the cytochrome $c$ synthetases ($Cc$F and $Cc$sA) that are discussed in this review. A CCHL structure, with and without substrates, will be very interesting.

It would be invaluable to reconstitute in vitro the holoCcmE release in step 1 and the step 2 synthetase reaction for system I and purified system II (CcsBA) and III (CCHL) synthetases. We need to learn much about the synthetase mechanisms for all three systems. Understanding the details of heme movement into or through the membrane, through channels or by other mechanisms, has been discussed above. Finally, more ecological and evolutionary studies are needed for further speculation on the advantages of each system.

**ACKNOWLEDGMENTS**

We thank Brian San Francisco for analysis of the natural distribution of systems, Kevin Moeller for discussions on imidazole reactivities, and Patrick Lane for illustrations. R.G.K. was supported by NIH grant GM 47909.

**REFERENCES**


heme lyase. CcmF forms a complex with the heme chaperone CcmE and CcmH but not with apocytochrome c. J. Biol. Chem. 277:7657–7663.


