Published in final edited form as:

J Mol Biol. 2018 April 13; 430(8): 1065–1080. doi:10.1016/j.jmb.2018.01.022.

Structurally Mapping Endogenous Heme in the CcmCDE Membrane Complex for Cytochrome *c* Biogenesis

Molly C. Sutherland^a, Joshua M. Jarodsky^a, Sergey Ovchinnikov^{b,c}, David Baker^{b,d}, and Robert G. Kranz^a

^aDepartment of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA

^bDepartment of Biochemistry, University of Washington, Seattle, Washington 98105, USA

^cPresent address: Harvard University, Cambridge, MA 02138, USA

^dHoward Hughes Medical Institute, University of Washington, Seattle, Washington 98105, USA

Abstract

Although many putative heme transporters have been discovered, it has been challenging to prove that these proteins are directly involved with heme trafficking *in vivo* and to identify their heme binding domains. The prokaryotic pathways for cytochrome *c* biogenesis, Systems I and II, transport heme from inside the cell to outside for stereochemical attachment to cytochrome *c*, making them excellent models to study heme trafficking. System I is composed of eight integral membrane proteins (CcmA–H) and is proposed to transport heme via CcmC to an external “WWD” domain for presentation to the membrane-tethered heme chaperone, CcmE. Herein, we develop a new cysteine/heme crosslinking approach to trap and map endogenous heme in CcmC (WWD domain) and CcmE (defining “2-vinyl” and “4-vinyl” pockets for heme). Crosslinking occurs when either of the two vinyl groups of heme localize near a thiol of an engineered cysteine residue. Double crosslinking, whereby both vinyls crosslink to two engineered cysteines facilitated a more detailed structural mapping of the heme binding sites, including stereospecificity. Using heme crosslinking results, heme ligand identification, and genomic coevolution data, we model the structure of the CcmCDE complex, including the WWD heme binding domain. We conclude CcmC trafficks heme via its WWD domain and propose the structural basis for stereochemical attachment of heme.

Graphical abstract

Corresponding Author: Robert G. Kranz, Campus Box 1137, One Brookings Drive, St. Louis, MO 63130, Phone: +1 (314) 935-4278, kranz@wustl.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Contributions: M.C.S. and R.G.K. designed research; M.C.S. and J.M.J. performed research; S.O. and D.B. predicted structures shown in Figs. 8, S6, S7, S8, data acquisition and interpretation; M.C.S., J.M.J., and R.G.K. analyzed data; M.C.S. and R.G.K. wrote the paper.

Conflict of Interest: The authors declare no conflicts of interest.

Keywords

cytochrome *c* biogenesis; heme; heme trafficking; heme binding site; CcmC structure

Introduction

Heme is synthesized in the cytoplasm of prokaryotes, whereas in eukaryotes the final step of synthesis, iron insertion by ferrochelatase, occurs inside the mitochondrion (or chloroplast). Because labile (free) heme is toxic and yet is required by many proteins essential for cellular function like hemoglobins and cytochromes, heme must be specifically trafficked throughout the cell [1–4]. Candidate heme chaperones and membrane transporters have been proposed based largely on genetic studies. For example, FLVCR1 and FLVCR2 are members of the major facilitator superfamily of membrane proteins, implicated in heme export for erythropoiesis and in anemia [5–7]. The HRG1 family with four transmembrane domains (TMDs) is involved in heme uptake by the heme auxotroph *C. elegans* [8,9], and in human heme transport [10]. Heme transporters of the ATP-binding cassette (ABC) family have also been reported in eukaryotes and bacteria (reviewed in [3]). Since the structures and biochemistry of these putative heme traffickers are poorly understood, the approach to prove heme transport has been to examine phenotypes, such as anemia. Recently, heme sensor proteins have been developed that allow *in vivo* monitoring of the labile heme pool in various compartments in the cell [11–14]. Nevertheless, controversy remains on whether heme is the substrate for transporters such as FLVCR proteins [15,16]. Here we develop an approach to trap and map endogenous heme binding within putative transporters and chaperones.

The pathways for prokaryotic cytochrome *c* biogenesis called Systems I and II have been excellent models for heme transport, chaperoning, heme modification, and redox chemistries [17–23]. However, many outstanding questions relate to binding sites of heme during trafficking and attachment to cytochrome *c* (Fig. 1A). In System I, composed of eight integral membrane proteins (CcmA–H), the first step in biogenesis involves putative heme transport through CcmC via the CcmABCD complex [24–26], to the periplasmic heme chaperone, CcmE (Fig. 1A) [27,28]. Release of holoCcmE (i.e. with heme) requires the full CcmABCD complex, an ABC transporter release complex. In the second step, holoCcmE transports heme to the holocytochrome *c* synthetase, CcmF/H, for attachment to apocytochrome *c* [29,30]. Here, we focus on step 1 of the System I pathway, specifically on the transfer of heme from CcmC to CcmE. A structure for apoCcmE* (no TMD) has been determined by NMR (Fig. 2A–C) [31], but not for holoCcmE, thus it is unclear where heme binds after or during its delivery by the CcmC protein. Two orientations of CcmE have been

suggested for the binding to heme (Fig. 2B, C) [26] but the heme binding site remains unknown [32]. The CcmC protein, comprised of six TMDs, is necessary for heme delivery to CcmE [28,33,24–26]. The small CcmD protein with a single transmembrane domain interacts CcmC and is required for holoCcmE release [34]. CcmC was originally shown to have a conserved periplasmic WWD domain (Fig. 1A) [35], a sequence motif also present in CcmF (the cytochrome *c* synthetase of System I), and CcsBA (also called ResBC [21]), the System II cytochrome *c* synthetase (Fig. 1B) [35,36]. This domain (Fig. 1B) is the defining sequence feature of a superfamily of transmembrane proteins sometimes referred to as the “putative heme handling proteins (HHPs)” [37], yet there is no evidence to date of direct binding of heme to this domain. In Systems I and II, a major working hypothesis has been that the WWD domain binds heme, presenting the two vinyl groups of heme (Fig. 1C) to the cognate acceptor protein (CcmE for CcmC, apocytochrome *c* for CcmF and CcsBA, see Fig. 1A) [17,33,38].

Prior studies using structurally known soluble cytochromes *b* [39–41] have shown that substituting a cysteine for a residue near a heme vinyl group results in spontaneous thioether attachments (Fig. 1D), here called a cysteine/heme “crosslink”. We reasoned that the chemical reactivity of either the 2- or 4-vinyl groups of heme towards a cysteine thiol might ‘trap’ endogenous heme and thus be useful in mapping heme binding sites, even those that are transient in nature and in membrane proteins. Engineering 16 cysteine substitutions in CcmE and 17 in the WWD domain of CcmC, we demonstrate in each case that heme is crosslinked to specific cysteine variants. To test the structural models, double crosslinks to both vinyl groups of heme were engineered, which further facilitated stereochemical structural mapping of the heme binding sites. We provide the first conclusive evidence the WWD domain directly binds heme and that its function is to position the vinyl groups of heme to its cognate acceptor.

Results

Mapping the heme binding site in CcmE

An intermediate CcmCDE membrane complex that contains heme was purified, made possible by inactivating CcmAB, thus preventing holoCcmE release from the complex (see Fig. 2E) [25,26]. To determine how heme is trafficked by this complex, we set out to structurally map where heme binds in both CcmE and CcmC in this complex. Using the structure of apoCcmE* [31], 16 residues mostly conserved (red) or semi-conserved (cyan) were substituted with cysteine in full-length CcmE (Fig. 2A), on either of two surface orientations displayed in Fig. 2B, C. The CcmE cysteine substitutions were engineered in the H130A background. Many studies on holoCcmE* by the Thöny-Meyer and Ferguson groups [27,31,32,42–45] have shown that His130 is the residue in CcmE that normally has a covalent link to a vinyl group of heme, resulting in retention of heme in the CcmE polypeptide upon SDS-PAGE (Fig. 2E, lane 4). A resonance Raman study suggests that the covalent link is formed to the 2-vinyl of heme [45], thus we assume CcmE covalently binds to the 2-vinyl of heme for purposes of discussion here. Unfortunately, the location of His130 in apoCcmE* has not facilitated understanding which CcmE orientation or surface heme is located (Fig. 2B or C). In the CcmCDE(H130A) complex, heme is present as *b* heme (Fig.

1C), which migrates as “free” heme at the SDS-PAGE dye front (Fig. 2E, lane 5 and Supplemental Fig. S1). After purifying each CcmCDE cysteine variant complex; UV-vis spectra, SDS-PAGE/Coomassie gels, heme stains, and CcmE westerns were performed (Supplemental Fig. S1, 2, Supplemental Table S1). Of the 16 complexes, only CcmE (H130A/E132C) possessed a covalent crosslink to heme (Fig 2E, lane 6; Supplemental Fig. S2F). CcmCDE(WT) exhibited UV-vis spectra with a unique split α -peak (552 nm and 559 nm) upon reduction (Fig. 2F); CcmCDE(H130A) and the CcmCDE(H130A/E132C) showed similar UV-vis absorptions to each other and similar levels of heme were present in all three complexes (Fig. 2F–H). A reduced pyridine hemochrome spectra (that reflects the number of covalent bonds to the heme molecule) of each complex indicated that the α -peak (Fig. 2I) of CcmCDE(H130A/E132C) was in between CcmCDE(WT) and CcmCDE(H130A). This is consistent with a single covalent linkage of the E132C to a heme vinyl group, as established with SDS-PAGE (Fig. 2E, lane 6 and Supplemental Fig. S2F).

Four of the CcmE cysteine substitutions (F37, G64, F103, G113) resulted in unstable proteins, likely due to folding defects, thus very little of these CcmE polypeptides were present in the complex (Supplemental Fig. S2A–E, Supplemental Table S1). Some of the cysteine substitutions were tested in the wildtype CcmE (with His130) background for function, i.e. the ability to attach heme to cytochrome *c* with a complete System I pathway (Supplemental Fig. S3, Table S1). Consistent with the folding defect, CcmE(G64C) showed reduced function (15%). The CcmE(E132C) variant had approximately 60% function, suggesting that His130 is still capable of forming its attachment (see ‘double-crosslinking’ below). The E132C cysteine/heme crosslink was used as a guide to model heme vinyl groups into the surface cavities of apoCcmE* (Fig. 2B–D). For either orientation (Fig. 2B or C), the E132C thiol was close enough (to His130) to form a thioether to the 2-vinyl. However, in the orientation in Fig. 2C, a conserved cavity (red, labelled the “4-vinyl pocket”) is 10 Å from the 2-vinyl pocket, defined by His130 and E132C. The 4-vinyl pocket is formed from conserved residues involved in folding (e.g. Phe37, Gly64). In heme, the 2-vinyl is 10 Å from the 4-vinyl and can be modeled into the pockets (Fig. 2D). Further evidence for designating this the 4-vinyl pocket is presented below in the double-crosslinking experiments, where it is also confirmed that E132C crosslinks to the 2-vinyl group of heme.

The WWD domain of CcmC crosslinks to heme

To address the role of CcmC in delivery of heme to CcmE, the cysteine/heme crosslinking approach was applied to the conserved WWD domain of CcmC (Fig. 3A). CcmC is proposed to traffick heme from the cytoplasm to the WWD domain where heme is liganded by two conserved histidine residues (P-His1, 2) and its vinyl groups positioned for interaction with CcmE (Fig. 3A) [25,26,28]. Cysteine substitutions were engineered in CcmC in the CcmE(H130A) background and each of the 17 CcmCDE(H130A) complexes were initially assayed for stability (Supplemental Fig. S4A, Fig. 3C) and crosslinked heme (Supplemental Fig. S4B, Fig. 3B, D). If a CcmC WWD cysteine variant crosslinked to heme, an increase is observed in heme signal in the CcmC polypeptide upon SDS-PAGE (Fig. 3B, D). Six variants (asterisks* in Fig. 3B) were identified with a CcmC to free heme ratio greater than 2X CcmCDE(H130A) and were selected for further analysis. Because

CcmC(G115C)DE(H130A), CcmC(W123C)DE(H130A) and CcmC(W125C)DE(H130A) were unstable in some preparations, only CcmC(W114C)DE(H130A), CcmC(D126C)DE(H130A) and CcmC(R128C)DE(H130A) were further characterized. These variants were stable (Fig. 3C), crosslinked heme at CcmC and had reduced levels of co-purified *b*-heme (Fig. 3D), suggesting that a cysteine/heme crosslink was formed. Each of the three CcmC variants were tested for function by engineering the cysteine substitutions into the wild type CcmE background. Each substitution resulted in at least a 75% reduction in function, as determined by levels of cytochrome *c*₄ assembled (Fig. 3E, F).

Given their low function in the pathway, we further probed the heme that is present in each complex and the ability to form the CcmCDE complex. Total heme in the cysteine/heme crosslinked complexes were determined by Soret absorbance (~410 nm) of UV-vis spectra (Fig. 4A, B). The variants co-purified with ~25% of heme compared to CcmCDE(H130A). Previous results suggested that heme is required for the stable interaction of CcmC and CcmE ([26], see discussion). Consistent with this, CcmE co-purified at lower levels with these variants compared to CcmCDE(H130A) (Fig. 4C, D). The cysteine/heme crosslinking variants have similar UV-vis spectral profiles to CcmCDE(H130A) (Fig. 4E–G). Pyridine hemochrome spectra (Fig. 4H) showed that CcmC cysteine/heme crosslinking variants displayed a range of reduced α -peaks from 553–555 nm. This result indicates CcmC(W114C)DE(H130A) has a single crosslinked heme, while CcmC(D126C)DE(H130A) and CcmC(R128C)DE(H130A) are mixtures of crosslinked and *b*-type heme, consistent with the heme stain upon SDS-PAGE from these complexes (Fig. 3D).

Orientation of heme in the WWD domain of the CcmC-CcmE complex, as determined by double-crosslinking of heme to both CcmE and CcmC

If heme is stereochemically oriented in the WWD domain, CcmC cysteine variants should form the crosslink to a specific heme vinyl group (2- or 4-vinyls). For example, CcmC(W114C) might crosslink to a different heme vinyl than CcmC(D126C) or CcmC(R128C). To examine stereochemistry, these cysteine substitutions in CcmC were engineered in the WT CcmE background (containing His130). A previous study suggests CcmE forms a covalent adduct to heme 2-vinyl via His130 [45]. We tested if the CcmC cysteine variants formed the CcmE His130 adduct (Fig. 5A(1)). Three possible crosslinks/adducts are diagrammed in Fig. 5A(2–4). The double crosslink (Fig 5A(2)) possibility would define which cysteine substitution in the WWD domain crosslinks to the 4-vinyl. In this case, a double crosslink would manifest from a single heme covalently attached to both CcmC and CcmE. GST affinity purification of the CcmC(W114C)DE(WT) variant resulted in a new high molecular weight band of ~72 kDa upon SDS-PAGE (Fig. 5B, lane 3), corresponding to the size of double crosslinked CcmC/heme/E(WT) co-complex. This variant displayed a unique heme stain pattern with crosslinked heme at the CcmE polypeptide, the CcmC polypeptide and the CcmC/heme/E(WT) double crosslinked complex (Fig. 5C, lane 3). In contrast, CcmC(D126C)DE(WT) and CcmC(R128C)DE(WT) heme stained only the CcmE polypeptide, similar to CcmCDE(WT), demonstrating that these variants formed only the His130 adduct in CcmE (Fig. 5C, lanes 1, 4 and 5). The 72 kDa polypeptide in CcmC(W114C)DE(WT) was a double crosslinked CcmC/heme/E(WT) co-

complex, as determined by western blotting with antisera to CcmE (Fig. 5D, lane 3) and GST (Fig. 5E, lane 3). An additional double crosslink was identified in CcmC(W123C)DE(WT) (Fig. S5). These results indicate that CcmC(W114C) and CcmC(W123C) form a crosslink to the heme 4-vinyl because the CcmE(H130) heme adduct is assumed to be to the 2-vinyl [27,44,45].

Because CcmC(D126C)DE(WT) and CcmC(R128C)DE(WT) did not form double crosslinked CcmC/heme/E(WT), we conclude that these cysteines crosslink to heme 2-vinyl only in the absence His130 (i.e. in CcmE(H130A)). In the presence of CcmE(H130), it is likely that a faster rate of formation occurs for the natural CcmE(H130) adduct (see Fig. 5A(4)). To confirm this, variants were further characterized by UV-vis spectra (Fig. 5F–H). CcmC(D126C)DE(WT) has a spectral profile similar to CcmCDE(WT) (Fig. 2F) with a diagnostic split α -peak (Fig. 5G), confirming that in this complex nearly all heme is bound to CcmE(WT) via the His130 adduct. In contrast, CcmC(W114C)DE(WT) has a single symmetrical α -peak at 555 nm (Fig. 5F), shifted from the CcmE(H130A) background (Fig. 4E), representing a mixture of covalent adducts, as shown by SDS-PAGE (Fig. 5C, lane 3). CcmC(R128C)DE(WT) shows an α -peak at 555 nm and a shoulder at 558 nm, suggesting a perturbation of the heme environment (Fig. 5H). Pyridine hemochrome spectra revealed that all variants have a reduced α -peak of 552 nm (Fig. 5I), demonstrating that all have a single covalent bond to heme and that in the case of CcmC(W114C)DE(WT) much of the heme is attached to CcmE or CcmC (Fig. 5I, blue), as suggested by the heme stain result (Fig. 5C, lane 3).

If CcmC(W114C) is crosslinking to the 4-vinyl and CcmE(E132C) to the 2-vinyl, a double crosslink should also result with this pair of substitutions (Fig. 6A). A double crosslink would additionally confirm that the CcmE(E132C) is formed to the 2-vinyl (like His130) and not 4-vinyl group. We engineered CcmC(W114C) in the CcmE (H130A/E132C) background and tested for crosslinking. Fig. 6B (lane 2) shows that purified CcmC(W114C)DE(H130A/E132C) complex yielded a 72 kDa polypeptide upon SDS-PAGE. The 72 kDa polypeptide stained for heme (Fig 6C, lane 2), and it reacted with antisera to CcmE (Fig. 6D, lane 2) and GST (Fig. 6E, lane 2), thus a intermolecular double crosslink was formed (between CcmE and CcmC, tethered covalently by heme). We conclude that CcmC Trp114 is adjacent to the 4-vinyl and CcmE His130 and Glu132 to the 2-vinyl in the CcmCDE complex. We attempted to generate an intramolecular CcmC double crosslink to both vinyl groups by combining two WWD domain cysteine substitutions, W114C with either D126C or R128C (in a CcmE(H130A) background). Unfortunately, such double CcmC mutants only bound a small amount of heme and CcmE (Fig 6B–D, lanes 3–4), thus we could not test for double crosslinking, although this result underscores the importance of WWD residues in binding heme.

Discussion

Using a cysteine/heme crosslinking approach on the CcmCDE membrane complex, the orientation of heme binding in CcmE was determined, as well as proof that heme directly binds to the WWD domain of CcmC with a specific orientation. These structural mapping results, combined with recent structural modeling advancements for membrane proteins

[46,47], facilitate discussion of the mechanisms underlying holoCcmE formation (Step 1 in Fig. 1A). We propose below the structural basis of the heme trafficking mechanism leading to stereochemical attachment of heme to CcmE (for CcmCDE). For holoCcmE this trafficking culminates in the covalent attachment of the 2-vinyl of heme to the CcmE His130 residue. For all cytochromes *c* (CXXCH motifs), trafficking ends at the cytochrome *c* synthetase (CcmF and CcsBA), resulting in thioether formation between the first cysteine to 2-vinyl and the second cysteine to 4-vinyl of heme. For the Systems I and II biogenesis pathways, a key domain proposed to be involved in the CXXCH stereochemical attachment is the WWD domain (in CcmF or CcsA). Our study with the CcmC WWD domain offers insight into this superfamily (CcmC, CcmF, CcsA).

Structure and formation of the CcmCD complex and interaction with CcmE

CcmC has six transmembrane domains, and has been previously shown to be part of a CcmCD complex [33,34]. CcmD is a small 62 residue protein with a single TMD and N-terminus facing the periplasmic space [34,48]. Although CcmD is not essential for heme attachment to CcmE [28,34], it is required for release of holoCcmE from the CcmCDE complex, which occurs upon binding of CcmAB and ATP hydrolysis [24,34]. Rosetta *de novo* structure programs using genomic coevolution data predict the CcmC structure shown in Fig. 7A and B. Further analyses of coevolved residues between CcmC and CcmD allows modeling of CcmD in the CcmCD complex (Fig. 7A and B). CcmD Tyr17 is in the conserved “GlyXxxTyr” N-terminal motif [34] and is predicted to interact with the WWD domain (Met 118) of CcmC (blue residues in Fig. 7A and B). Four TMDs in CcmC are predicted to interact with each other (TM2–5) as well as CcmD in the structure (Fig. 7B).

Heme does not stably bind to CcmCD. That is, endogenous heme does not co-purify with the CcmCD complex without CcmE [26]. Similarly, CcmE does not bind stably to CcmCD without heme [26]. Thus, heme is a central component required for formation and purification of the holo(heme)CcmCDE complex [26]. We propose that CcmE first transiently interacts with CcmCD, exposing the heme binding WWD domain and that subsequent heme binding facilitates stable CcmCDE complex formation (see Fig. 1A). The initial CcmCD:CcmE interaction likely requires specific residues in each. For example, the CcmE(K129C/H130A) variant exhibits some distinct properties consistent with such an interaction, with less heme and CcmE bound in the complex and a reduction in function for the CcmE(K129C) variant (Fig. S1, S3, Table S1). It is possible that CcmE Lys129 is involved in initial interaction with CcmC, prior to heme binding, in keeping with its key location (see K129 in Fig. 2C). Here we confirm with many CcmC and CcmE variants that the levels of CcmE in complexes are always proportional to the level of heme. In the absence of CcmAB, heme in these ternary complexes is tightly bound, consistent with a requirement for the CcmAB-mediated release of holoCcmE from the complex (Fig. 1A, see “holoCcmE release” below).

Heme binding in the CcmCDE complex

Because no structure for CcmCD bound or released holoCcmE is known, we used heme/cysteine crosslinking to help define where vinyl groups of heme reside in the CcmCDE complex and holoCcmE. CcmE(E132C) is crosslinked to the 2-vinyl group of heme in a

H130A variant (Fig. 6B–E). This information facilitated modeling of heme in CcmE revealing the “2-vinyl” and “4-vinyl” pockets (Fig. 2C, D). We then performed extensive crosslinking studies on CcmC, for the first time proving that the conserved WWD domain directly binds heme in System I (CcmC) (Fig. 3). CcmC trafficks heme to the WWD domain where its axial ligands to heme iron are periplasmic histidines (His60 and His184) [26]. For CcmC, the mechanism to traffick heme to the WWD domain is likely from the cytoplasm through the transmembrane domains although this has not been proven (Fig. 3A) [26]. Double-crosslinking studies (Figs. 5, 6) successfully demonstrated that CcmC Trp114 crosslinks to the 4-vinyl of heme. We propose that heme in the WWD domain is oriented with 4-vinyl and 2-vinyl such that interaction with the acceptor protein is stereospecific. As structurally described below, this stereochemically aligns heme with the proposed 2- and 4-vinyl pockets of CcmE, allowing for adduct formation to His130 of CcmE (Fig. 2D). Although it is likely that 2-vinyl is attached to CcmE His130, even if it is the 4-vinyl and the heme reversed, the mapping and stereospecificities described here are still valid, including conclusions from the double-crosslinking experiments.

Ovchinnikov *et al* (2017) note that for the hundreds of predicted membrane protein structures, a limitation of modeling includes ligand or co-factor binding domains as well as segments joined by TMDs [47]. For the six TMD CcmC, such inaccuracies are therefore expected for the WWD domain and in adjoining periplasmic domains that contain the two histidines. We have modeled the WWD and adjoining domains (Fig. 8A–C) using the following parameters from our experimental studies as described in the materials and methods: i) the WWD domain forms a heme binding site with vinyl groups exposed; ii) Trp114 and Gly115 are within 2 Å of the 4-vinyl group; iii) Trp123 is near the 4-vinyl group; iv) P-His1, His60 (in periplasmic loop 1), is an axial ligand to the heme iron; v) P-His2, His184 (in periplasmic loop 3), is the second axial ligand on the opposite plane of heme (to axial ligand P-His1); vi) the conserved tryptophans that define the WWD domain (W119, W123, W125) interact directly with the heme. Using these constraints and low-energy modeling criteria, a heme binding site is predicted in Fig. 8A–C. An important feature is that the 4-vinyl and 2-vinyl groups are surface exposed to “present” to the acceptor proteins, in this case CcmE. The CcmCDE complex with structurally modeled heme is shown in Fig. 8D and E, where the 2- and 4-vinyl groups can be inserted into their respective pockets in CcmE. The CcmE His130 imidazole is shown in Fig. 8E adjacent to the exposed 2-vinyl. The WWD heme binding domain predicted here is a working model based in part on known functions and experimental findings, as detailed in Methods. The model is for illustrative purposes and will be useful for testing hypotheses on structure/function of superfamily members. Supplemental Fig. S7 displays the ensembles of heme-binding sites (WWD domain) upon which the model was derived.

HoloCcmE release from the CcmCDE complex

Experimental results indicate a very stable holoCcmCDE complex, requiring CcmAB binding and ATP hydrolysis (via the CcmA ATP-binding cassette) for release of holoCcmE [24]. This process involves release of the two axial histidine ligands (in CcmC), with replacement of one by Tyr134 of CcmE [32,42,45] (Fig. 2C). We speculate that ATP hydrolysis by CcmA induces a conformational change in CcmCD that alters heme binding in

the WWD domain. Given the specific co-evolved interactions between CcmD and the CcmC WWD domain (Fig. 7A, B) and the requirement of CcmD for release, we propose a working model for heme binding. In this proposal, CcmAB complexed with holoCcmCDE, undergoes ATP hydrolysis (by CcmA). This hydrolysis disrupts heme:WWD interactions and favors CcmD contacts with the WWD domain (e.g. CcmD Tyr17 to CcmC Met118), thereby releasing heme (as holoCcmE). Next, binding of apoCcmE to CcmCD induces the WWD:heme binding conformation, which then restricts the CcmD:WWD domain interaction. Each ATP hydrolysis and release begins the cycle again. Once released, CcmE would naturally fold around the heme, likely maintaining the 2-vinyl and 4-vinyl heme interactions. This is analogous to release of holocytochromes *c* from their holocytochrome *c* synthetases, where it is generally agreed that complete folding occurs after heme attachment [18,21–23,49].

Relevance of heme/cysteine crosslinking to cyt *c* biogenesis pathways (CXXCH thioether attachments) and other putative heme trafficking proteins

Our data demonstrates that heme resides in the CcmC WWD domain for stereochemical positioning and subsequent attachment to apoCcmE. Because the conserved WWD domain is also present in the holocytochrome *c* synthetases, CcmF (System I) and CcsBA (System II) (Fig. 1B), these likely function in an analogous manner to position and attach heme to the apocytochrome *c*. Ongoing analysis of the CcsBA WWD domain also demonstrates specific heme-cysteine crosslinks and supports this model, enabling structural and mechanistic comparisons between System I and System II. Previous studies have shown that alanine substitutions in selected residues of the WWD domain result in a reduced ability to mature cytochrome *c* [26,33,38,50,51]. The cysteine/heme crosslinking studies here have begun to elucidate the molecular basis for these defects, making it likely that the alanine substitutions alter heme binding or positioning in the WWD domain.

For CcmF, heme from CcmE is the substrate for attachment to cytochrome *c* [29], whereas in CcsBA heme is proposed to be transported through the transmembrane regions to the WWD domain [51,52]. The structural predictions in Fig. 8A–C on the CcmC WWD domain are relevant to CcsBA (and CcmF), likewise suggesting the mechanism for stereochemical attachment to the apocytochrome *c* (CXXCH) acceptors, as long as the CXXCH motifs are themselves oriented upon binding (analogous to CcmE orientation).

The structures proposed here and heme crosslinks discovered will be useful in testing the dynamic processes of heme trafficking and the chemistries involved in synthetase reactions. Future cryo-EM and other structural elucidations of these heme traffickers and intermediates will also be informed by our structural models. This approach of cysteine/heme crosslinking may be useful in mapping endogenous heme binding sites in other putative heme transporters and in substantiating their roles in heme trafficking.

Materials and Methods

Bacterial Growth Conditions

Escherichia coli strains were grown in Luria-Bertani broth (LB, Difco) at 37°C and 200 rpm with appropriate inducing reagents and selective antibiotics: Isopropyl β -D-1-thiogalactopyranoside (IPTG, Gold Biotechnology), 1 mM or 0.1 mM; arabinose (Alfa Aesar), 0.2% (w/v); carbenicillin, 50 μ g/mL; chloramphenicol, 20 μ g/mL.

Construction of cysteine variants

All cloning was performed in *E. coli* NEB 5- α and XL1-Blue. Cysteine substitutions were engineered using the QuikChange II Site-directed Mutagenesis Kit (Agilent Technologies) according to the manufacturer's instructions. Substitutions were confirmed by sequencing. A complete list of primers and templates can be found in Table S2.

Protein Purification

Affinity purifications of GST:CcmC fusions were performed as described in [53]. Briefly, *E. coli* strain RK103 was used for protein expression. 10 mL starter cultures were diluted 1:100 into 1L LB with appropriate antibiotics and grown at 37°C and 200 rpm to an OD₆₀₀ of ~1.2. Cultures were induced with 1 mM IPTG, grown for an additional ~16 hours, harvested by centrifugation and cell pellets stored at -80°C. Cells were resuspended in GST buffer (4.3 mM Na₂HPO₄, 1.5 mM KH₂PO₄, 2.7 mM KCl, 140 mM NaCl, pH7.3), supplemented with 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma-Aldrich) and 1 mg/mL egg white lysozyme (Sigma-Aldrich). Cells were lysed by sonication (Branson250 sonicator), cleared of cell debris by centrifugation at 24000g for 30 minutes at 4°C, followed by separation of soluble and membrane fractions via high-speed ultracentrifugation at 100000g for 45 minutes at 4°C. Membrane pellets were solubilized in GST buffer with 1% *n*-dodecyl- β -D-maltopyranoside (DDM; Anatrace) and affinity purified by the batch method with glutathione agarose (Pierce). Columns were washed by gravity flow, eluted in 4 mL GST buffer supplemented with 0.02% DDM and 20 mM L-glutathione (Sigma-Aldrich) and concentrated in a 30 kDa Amicon filter. Note: no exogenous heme was added to cells or protein preparations. Heme that co-purifies with the affinity purifications is endogenous and interaction with proteins occurred during growth and induction in *E. coli*. Protein concentrations were determined by Bradford Assay (Sigma-Aldrich).

I with the following modifications: 20 mM Tris-pH8, 100 mM NaCl buffer used.

CcmE Modeling

PyMOL was used to visualize PDB 1SR3 of apoCcmE*, as determined by Enggist *et al* [31]. Using the orientations discussed in [26], "vinyl pockets" were measured using PyMOL software. Heme (HEMsdf) was introduced into the apoCcmE structure and manually docked into the 2-vinyl and 4-vinyl pockets. CcmE residue Lys129 was omitted from CcmE when heme was docked (Fig. 2D), to better visualize heme docking.

Heme stains, Immunoblots and Quantification

Heme stains were performed as previously described by [54] with the following modifications. Proteins were transferred to 0.2 µm Immobilon P^{SQ} (Millipore). Two membranes were layered to capture free heme that transferred through the first membrane. Imaging was performed with a LAS-1000 Plus (Fujifilm-GE Healthcare) or LI-COR Odyssey Fc (LI-COR Biosciences). Heme abundance was determined by densitometry performed with ImageJ [55] or Image Studio Lite Version 5.2 software (LI-COR Biosciences). Note that the total free heme value was determined by analysis of heme on both membranes. For immunoblots, proteins were separated via SDS-PAGE, transferred to 0.45 µm Amersham Protran nitrocellulose (GE Healthcare Life Sciences) and probed with the following antibodies: CcmE – α-CcmE at 1:15000 [24], GST:CcmC – α-GST at 1:5000 (Sigma-Aldrich). Protein A peroxidase (Sigma-Aldrich) and developed with Immobilon Western Chemiluminescent HRP substrate (Millipore) or IRDye 800 CW (LI-COR Biosciences) were used as secondary labels. Imaging and quantitation was performed as above and total proteins stains were performed with Coomassie Blue or SYPRO Ruby Protein. Note that for ease of viewing, Coomassie stains are false colored blue, heme stains false colored red and western blots false colored green.

Functional Assays

System I, CcmE cysteine variant function – The appropriate variant was co-expressed with pUCA6P4-Comp CcmAB (pRGK365) and pBAD CcmF:HisGH (pRGK388) in RK111 (RK103 with cyt *c4*:His chromosomal integrate). System I, CcmC cysteine variant function - The appropriate variant was co-expressed with cytochrome *c4* (pMCS85) in RK103. 1 mL starter cultures were grown overnight in LB broth supplemented with appropriate antibiotics at 37°C with shaking. Cultures were diluted 1:5 into LB with appropriate antibiotics and grown at 37°C and 200 rpm for 3 hours. Proteins were induced with 0.1 mM IPTG (System I or System II plasmids) and 0.2% arabinose (cytochrome *c4*) and grown for 3 additional hours, harvested by centrifugation for 10 minutes at 5000g and stored at –80°C. Cells were lysed with 200 µL B-per reagent (Thermo-Scientific) according to the manufacturer's instructions. Protein concentrations were determined with a Nanodrop 1000 spectrophotometer (Thermo-Scientific) and 100 µg of protein was separated by SDS-PAGE, analyzed for cytochrome *c* maturation by heme stain and quantified by densitometry with ImageJ [55], values were normalized to the wildtype System I plasmid.

UV-visible absorption spectroscopy

UV-vis absorption spectra were collected aerobically at room temperature with a Shimadzu UV-1800 spectrophotometer using UVProbe 2.43 software. Spectra were obtained in the elution buffer used for affinity purification. Quantitation of total heme levels were determined using 50 µg of air-oxidized purified protein with collection from 360–800 nm. Absorbance of the Soret peaks were normalized to wildtype. Spectral characterization of “as purified” and reduced proteins were collected from 360–800 nm using a protein concentration that resulted in a minimum Soret absorbance of 0.2. Reduced spectra were generated chemically by addition of solid sodium dithionite (sodium hydrosulfite, Sigma-Aldrich) to the sample. Pyridine hemochrome spectra were performed as previously

described [56]. Briefly, a protein concentration that resulted in a Soret absorbance of at least 0.3 was used and 0.5 M NaOH and pyridine were added to the sample to final concentrations of 100 mM NaOH and 20% pyridine (v/v). Samples were chemically reduced as above and spectra were recorded from 500–600 nm.

CcmCDE Modeling

Alignment generation—Starting with CcmC, CcmD and CcmE *E. coli* sequences, Jackhmmer from the HMMER package version 3.1b1 [57] was used to search against the European Nucleotide Archive (ENA) database of protein sequences from release 132, to generate a multiple sequence alignment (MSA). To prevent paralogs from contaminating the alignment, different e-values were scanned until only one gene per genome was detected in the MSA. E-values of 1E-40, 1E-2 and 1E-40 were selected for CcmC, CcmD and CcmE, respectively, with 8 iterations each. Paired alignments: Using the genomic information from ENA, paired alignments were made for CcmC_CcmD, CcmC_CcmE and CcmD_CcmE. To prevent paralogs, only sequences from the same operon were used. To determine if the sequences were from the same operon, only sequences that were within five genes apart in the genome were considered.

Contact prediction—HHfilter was used to filter redundant sequence (HHsuite version 2.0.15; -id 90 -cov 75) [58] and positions that had more than 50% gaps were removed. The resulting alignment were inputted into GREMLIN v2.01 [59] for contact prediction.

Fragment picking—Given the discrepancy in secondary structure prediction between PSIPRED [60] and RaptorX [61] (see Supplemental Fig. S6), for fragment picking, both were used. Half the fragments came from PSIPRED and the other half from RaptorX. For fragment picking, Rosetta was used [62]. The structural profiles were disabled during picking, since these are biased towards soluble proteins. Fragment database from March 2015 was used from PDB30 dataset [47].

Model building—For the initial global sampling, Rosetta abinitio protocol in combination with GREMLIN restraints were used as described in [47] for trans-membrane proteins. In addition to these restraints, a bounded restraint was used between 0–6 angstroms of the CB-CB atoms of the two histidine residues (H60 and H184).

The CcmD and CcmC sequences were concatenated with a glycine linker as follows:

```
MGGYAFFVWLAVVMTVIPLVVLVHVSVMQHRAILRGV[GGGGGGGGGGGG]WFIPWLAIASVV
VLTVGWIWGFAPADYQQGNSYRIIYLHVPAAIWSMGIYASMAVAFIGLVQMKNANLAVA
AMAPIGAVFTFIALVTGSAWGKPMWGTWWVWDARLTSELVLLFLYVGVIALLWAFDDRRLAGR
AAGILVLIGVNLPIIHYSVEWWNTLHQGSTRMQQSIDPAMRSPLRWSIFGFLLLSATLTLMR
MRNLILL 20,000 models were generated. The top 10 models were compared for
convergence. The top model (selected by best rosetta+restraint energy) was then used
for further modeling.
```

Modeling in the heme—The glycine linker was removed along with the following loops: 41–63, 115–126 and 184–195. HMY (heme-like molecule), with the vinyl and propionates

replaced by methyl groups, was generated. The hybridization protocol in Rosetta [63] was used to sample the removed loops around HMY. HMY was allowed to rotate and translate. During the coarse-grained sampling, bounded restraints as described were used. During the full-atom refinement stage harmonic restraints were used. These include distance restraints between NE2 and FE1 atoms (mean: 2 angstroms, stdev 0.1), angle restraints between N1-FE1-NE2, N2-FE1-NE2, N3-FE1-NE2 and N3-FE1-NE2 (mean: 90 degrees, stdev: 10), and dihedral restraint between CG-CD2-NE2-FE1 (mean 180 degrees, stdev: 10). 5000 models were sampled. The heme placement converged in the top 10 models. These models were then screened for interaction with the WWD domain. More specifically, we looked for models where one of the methyl groups of HMY was near R128 and D126 and another methyl group was near W114, G115 and W123 (corresponding to locations where 2-vinyl and 4-vinyl would be). The only models that satisfied this condition buried the 2-vinyl site and exposed the propionates. Next, the models were screened so that the propionate faced down, 2-vinyl was exposed and 4-vinyl was interacting with W114, G115 and W123. Interestingly, in these models the R128 is near the propionate group, suggesting that the arginine may form ionic interaction and upon mutation to cysteine, would allow the heme to rotate and crosslink to 2-vinyl. Other explanations are possible, so we focus discussion on the double-crosslinked results (W114C (Fig. 5) and W123C (Fig. S5)), as well as conserved tryptophans.

The lowest energy model that satisfied the latter condition was selected for further sampling with the full heme. The WWD domain (114–128) was deleted to prevent bias from starting model. Ambiguous sigmoidal restraints (mean: 5 angstroms, slope: 4) were added between all available side-chain carbon atoms of W114, W119, W123, W125 and carbon atoms of heme, during both coarsened-grained and full atom mode of the hybridized protocol. This was to bias sampling towards interactions that maximize aromatic ring stacking between the tryptophan and heme. Besides these and the histidine restraints described before, additional restraints were added between R128 and propionate carbons, between W114, G115, W123, and 4-vinyl carbons. 1500 models were sampled. The top scoring model is shown in this paper. Due to the large number of degrees of freedom, the mutations involved in identifying the crosslinks, and uncertainties in modeling heme containing loops generally, the models in figure 8 are for illustrative purposes and hypothesis testing only. The models are not expected to be accurate at the level of the converged models in [47].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Cynthia L. Richard-Fogal and Alan Sariol for assistance with the CcmE cysteine variants; Joel A. Rankin and Ian M. Furey for assistance with the CcmC cysteine variants; Joel A. Rankin, Sarah Santiago, Deanna L. Mendez and the 2015 Washington University Summer Bio437 class for cloning assistance. This work was funded by the National Institutes of Health R01 GM47909 to R.G.K. and F32GM115020 to M.C.S.

References

1. Khan AA, Quigley JG. Control of intracellular heme levels: Heme transporters and heme oxygenases. *Biochim Biophys Acta BBA - Mol Cell Res.* 2011; 1813:668–682. DOI: 10.1016/j.bbamcr.2011.01.008
2. Hamza I, Dailey HA. One ring to rule them all: Trafficking of heme and heme synthesis intermediates in the metazoans. *Biochim Biophys Acta BBA - Mol Cell Res.* 2012; 1823:1617–1632. DOI: 10.1016/j.bbamcr.2012.04.009
3. Chiabrando D, Vinchi F, Fiorito V, Mercurio S, Tolosano E. Heme in pathophysiology: a matter of scavenging, metabolism and trafficking across cell membranes. *Front Pharmacol.* 2014; 5doi: 10.3389/fphar.2014.00061
4. Reddi AR, Hamza I. Heme Mobilization in Animals: A Metalloprotein's Journey. *Acc Chem Res.* 2016; 49:1104–1110. DOI: 10.1021/acs.accounts.5b00553 [PubMed: 27254265]
5. Quigley JG, Yang Z, Worthington MT, Phillips JD, Sabo KM, Sabath DE, Berg CL, Sassa S, Wood BL, Abkowitz JL. Identification of a Human Heme Exporter that Is Essential for Erythropoiesis. *Cell.* 2004; 118:757–766. DOI: 10.1016/j.cell.2004.08.014 [PubMed: 15369674]
6. Keel SB, Doty RT, Yang Z, Quigley JG, Chen J, Knoblauch S, Kingsley PD, De Domenico I, Vaughn MB, Kaplan J, Palis J, Abkowitz JL. A Heme Export Protein Is Required for Red Blood Cell Differentiation and Iron Homeostasis. *Science.* 2008; 319:825–828. DOI: 10.1126/science.1151133 [PubMed: 18258918]
7. Chiabrando D, Marro S, Mercurio S, Giorgi C, Petrillo S, Vinchi F, Fiorito V, Fagoonee S, Camporeale A, Turco E, Merlo GR, Silengo L, Altruda F, Pinton P, Tolosano E. The mitochondrial heme exporter FLVCR1b mediates erythroid differentiation. *J Clin Invest.* 2012; 122:4569–4579. DOI: 10.1172/JCI62422 [PubMed: 23187127]
8. Rao AU, Carta LK, Lesuisse E, Hamza I. Lack of heme synthesis in a free-living eukaryote. *Proc Natl Acad Sci.* 2005; 102:4270–4275. DOI: 10.1073/pnas.0500877102 [PubMed: 15767563]
9. Rajagopal A, Rao AU, Amigo J, Tian M, Upadhyay SK, Hall C, Uhm S, Mathew MK, Fleming MD, Paw BH, Krause M, Hamza I. Haem homeostasis is regulated by the conserved and concerted functions of HRG-1 proteins. *Nature.* 2008; 453:1127–1131. DOI: 10.1038/nature06934 [PubMed: 18418376]
10. White C, Yuan X, Schmidt P, Bresciani E, Samuel T, Campagna D, Hall C, Bishop K, Calicchio M, Lapierre A, Ward D, Liu P, Fleming M, Hamza I. HRG1 Is Essential for Heme Transport from the Phagolysosome of Macrophages during Erythrophagocytosis. *Cell Metab.* 2013; 17:261–270. DOI: 10.1016/j.cmet.2013.01.005 [PubMed: 23395172]
11. Song Y, Yang M, Wegner SV, Zhao J, Zhu R, Wu Y, He C, Chen PR. A Genetically Encoded FRET Sensor for Intracellular Heme. *Chem Biol.* 2015; 10:1610–1615. DOI: 10.1021/cb5009734
12. Hanna DA, Harvey RM, Martinez-Guzman O, Yuan X, Chandrasekharan B, Raju G, Outten FW, Hamza I, Reddi AR. Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors. *Proc Natl Acad Sci.* 2016; 113:7539–7544. DOI: 10.1073/pnas.1523802113 [PubMed: 27247412]
13. Yuan X, Rietzschel N, Kwon H, Walter Nuno AB, Hanna DA, Phillips JD, Raven EL, Reddi AR, Hamza I. Regulation of intracellular heme trafficking revealed by subcellular reporters. *Proc Natl Acad Sci.* 2016; 113:E5144–E5152. DOI: 10.1073/pnas.1609865113 [PubMed: 27528661]
14. Abshire JR, Rowlands CJ, Ganesan SM, So PTC, Niles JC. Quantification of labile heme in live malaria parasites using a genetically encoded biosensor. *Proc Natl Acad Sci.* 2017; 114:E2068–E2076. DOI: 10.1073/pnas.1615195114 [PubMed: 28242687]
15. Yuan X, Fleming MD, Hamza I. Heme transport and erythropoiesis. *Curr Opin Chem Biol.* 2013; 17:204–211. DOI: 10.1016/j.cbpa.2013.01.010 [PubMed: 23415705]
16. Ponka P, Sheftel AD, English AM, Scott Bohle D, Garcia-Santos D. Do Mammalian Cells Really Need to Export and Import Heme? *Trends Biochem Sci.* 2017; 42:395–406. DOI: 10.1016/j.tibs.2017.01.006 [PubMed: 28254242]
17. Kranz R, Lill R, Goldman B, Bonnard G, Merchant S. Molecular mechanisms of cytochrome *c* biogenesis: three distinct systems. *Mol Microbiol.* 1998; 29:383–396. [PubMed: 9720859]

18. Kranz RG, Richard-Fogal C, Taylor JS, Frawley ER. Cytochrome *c* biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. *Microbiol Mol Biol Rev.* 2009; 73:510–528. Table of Contents. DOI: 10.1128/MMBR.00001-09 [PubMed: 19721088]
19. Bonnard G, Corvest V, Meyer EH, Hamel PP. Redox Processes Controlling the Biogenesis of *c*-Type Cytochromes. *Antioxid Redox Signal.* 2010; 13:1385–1401. DOI: 10.1089/ars.2010.3161 [PubMed: 20214494]
20. Sanders C, Turkarslan S, Lee DW, Daldal F. Cytochrome *c* biogenesis: the Ccm system. *Trends Microbiol.* 2010; 18:266–274. DOI: 10.1016/j.tim.2010.03.006 [PubMed: 20382024]
21. Simon J, Hederstedt L. Composition and function of cytochrome *c* biogenesis System II. *FEBS J.* 2011; 278:4179–4188. DOI: 10.1111/j.1742-4658.2011.08374.x [PubMed: 21955752]
22. Stevens JM, Mavridou DAI, Hamer R, Kritsiligkou P, Goddard AD, Ferguson SJ. Cytochrome *c* biogenesis System I. *FEBS J.* 2011; 278:4170–4178. DOI: 10.1111/j.1742-4658.2011.08376.x [PubMed: 21958041]
23. Gabilly ST, Hamel PP. Maturation of Plastid *c*-type Cytochromes. *Front Plant Sci.* 2017; 8:1313.doi: 10.3389/fpls.2017.01313 [PubMed: 28798763]
24. Feissner RE, Richard-Fogal CL, Frawley ER, Kranz RG. ABC transporter-mediated release of a haem chaperone allows cytochrome *c* biogenesis. *Mol Microbiol.* 2006; 61:219–231. DOI: 10.1111/j.1365-2958.2006.05221.x [PubMed: 16824107]
25. Richard-Fogal CL, Frawley ER, Bonner ER, Zhu H, San Francisco B, Kranz RG. A conserved haem redox and trafficking pathway for cofactor attachment. *EMBO J.* 2009; 28:2349–2359. DOI: 10.1038/emboj.2009.189 [PubMed: 19629033]
26. Richard-Fogal C, Kranz RG. The CcmC:heme:CcmE complex in heme trafficking and cytochrome *c* biosynthesis. *J Mol Biol.* 2010; 401:350–362. DOI: 10.1016/j.jmb.2010.06.041 [PubMed: 20599545]
27. Schulz H, Hennecke H, Thöny-Meyer L. Prototype of a heme chaperone essential for cytochrome *c* maturation. *Science.* 1998; 281:1197–1200. [PubMed: 9712585]
28. Schulz H, Fabianek RA, Pellicoli EC, Hennecke H, Thöny-Meyer L. Heme transfer to the heme chaperone CcmE during cytochrome *c* maturation requires the CcmC protein, which may function independently of the ABC-transporter CcmAB. *Proc Natl Acad Sci.* 1999; 96:6462–6467. [PubMed: 10339610]
29. San Francisco B, Kranz RG. Interaction of holoCcmE with CcmF in heme trafficking and cytochrome *c* biosynthesis. *J Mol Biol.* 2014; 426:570–585. DOI: 10.1016/j.jmb.2013.10.025 [PubMed: 24513106]
30. San Francisco B, Sutherland MC, Kranz RG. The CcmFH complex is the system I holocytochrome *c* synthetase: engineering cytochrome *c* maturation independent of CcmABCDE. *Mol Microbiol.* 2014; 91:996–1008. DOI: 10.1111/mmi.12510 [PubMed: 24397552]
31. Enggist E, Thöny-Meyer L, Güntert P, Pervushin K. NMR structure of the heme chaperone CcmE reveals a novel functional motif. *Struct.* 1993; 10(2002):1551–1557.
32. Harvat EM, Redfield C, Stevens JM, Ferguson SJ. Probing the heme-binding site of the cytochrome *c* maturation protein CcmE. *Biochemistry.* 2009; 48:1820–1828. DOI: 10.1021/bi801609a [PubMed: 19178152]
33. Goldman BS, Beck DL, Monika EM, Kranz RG. Transmembrane heme delivery systems. *Proc Natl Acad Sci.* 1998; 95:5003–5008. [PubMed: 9560218]
34. Richard-Fogal CL, Frawley ER, Kranz RG. Topology and Function of CcmD in Cytochrome *c* Maturation. *J Bacteriol.* 2008; 190:3489–3493. DOI: 10.1128/JB.00146-08 [PubMed: 18326572]
35. Beckman DL, Trawick DR, Kranz RG. Bacterial cytochromes *c* biogenesis. *Genes Dev.* 1992; 6:268–283. [PubMed: 1310666]
36. Xie Z, Merchant S. The plastid-encoded *ccsA* gene is required for heme attachment to chloroplast *c*-type cytochromes. *J Biol Chem.* 1996; 271:4632–4639. [PubMed: 8617725]
37. Lee JH, Harvat EM, Stevens JM, Ferguson SJ, Saier MH. Evolutionary origins of members of a superfamily of integral membrane cytochrome *c* biogenesis proteins. *Biochim Biophys Acta BBA - Biomembr.* 2007; 1768:2164–2181. DOI: 10.1016/j.bbamem.2007.04.022

38. Hamel PP, Dreyfuss BW, Xie Z, Gabilly ST, Merchant S. Essential histidine and tryptophan residues in CcsA, a system II polytopic cytochrome *c* biogenesis protein. *J Biol Chem*. 2003; 278:2593–2603. DOI: 10.1074/jbc.M208651200 [PubMed: 12427766]
39. Barker PD, Ferrer JC, Mylrajan M, Loehr TM, Feng R, Konishi Y, Funk WD, MacGillivray RT, Mauk AG. Transmutation of a heme protein. *Proc Natl Acad Sci*. 1993; 90:6542–6546. [PubMed: 8341666]
40. Barker PD, Nerou EP, Freund SM, Fearnley IM. Conversion of cytochrome b562 to *c*-type cytochromes. *Biochemistry*. 1995; 34:15191–15203. [PubMed: 7578134]
41. Lin YW, Wang WH, Zhang Q, Lu HJ, Yang PY, Xie Y, Huang ZX, Wu HM. Converting Cytochrome *b5* into Cytochrome *c*-Like Protein. *ChemBioChem*. 2005; 6:1356–1359. DOI: 10.1002/cbic.200500030 [PubMed: 15977276]
42. Enggist E, Schneider MJ, Schulz H, Thöny-Meyer L. Biochemical and mutational characterization of the heme chaperone CcmE reveals a heme binding site. *J Bacteriol*. 2003; 185:175–183. [PubMed: 12486054]
43. Daltrop O, Stevens JM, Higham CW, Ferguson SJ. The CcmE protein of the *c*-type cytochrome biogenesis system: unusual *in vitro* heme incorporation into apo-CcmE and transfer from holo-CcmE to apocytochrome. *Proc Natl Acad Sci*. 2002; 99:9703–9708. DOI: 10.1073/pnas.152120699 [PubMed: 12119398]
44. Lee D, Pervushin K, Bischof D, Braun M, Thöny-Meyer L. Unusual heme-histidine bond in the active site of a chaperone. *J Am Chem Soc*. 2005; 127:3716–3717. DOI: 10.1021/ja044658e [PubMed: 15771504]
45. Uchida T, Stevens JM, Daltrop O, Harvat EM, Hong L, Ferguson SJ, Kitagawa T. The interaction of covalently bound heme with the cytochrome *c* maturation protein CcmE. *J Biol Chem*. 2004; 279:51981–51988. DOI: 10.1074/jbc.M408963200 [PubMed: 15465823]
46. Ovchinnikov S, Kinch L, Park H, Liao Y, Pei J, Kim DE, Kamisetty H, Grishin NV, Baker D. Large-scale determination of previously unsolved protein structures using evolutionary information. *eLife*. 2015; 4doi: 10.7554/eLife.09248
47. Ovchinnikov S, Park H, Varghese N, Huang PS, Pavlopoulos GA, Kim DE, Kamisetty H, Kyripides NC, Baker D. Protein structure determination using metagenome sequence data. *Science*. 2017; 355:294–298. DOI: 10.1126/science.aah4043 [PubMed: 28104891]
48. Goldman BS, Beckman DL, Bali A, Monika EM, Gabbert KK, Kranz RG. Molecular and immunological analysis of an ABC transporter complex required for cytochrome *c* biogenesis. *J Mol Biol*. 1997; 268:724–738. DOI: 10.1006/jmbi.1997.0992 [PubMed: 9175857]
49. Verissimo AF, Daldal F. Cytochrome *c* biogenesis System I: an intricate process catalyzed by a maturase supercomplex? *Biochim Biophys Acta*. 2014; 1837:989–998. DOI: 10.1016/j.bbabi.2014.03.003 [PubMed: 24631867]
50. Ahuja U, Thöny-Meyer L. Dynamic features of a heme delivery system for cytochrome *c* maturation. *J Biol Chem*. 2003; 278:52061–52070. DOI: 10.1074/jbc.M310077200 [PubMed: 14532274]
51. Frawley ER, Kranz RG. CcsBA is a cytochrome *c* synthetase that also functions in heme transport. *Proc Natl Acad Sci*. 2009; 106:10201–10206. DOI: 10.1073/pnas.0903132106 [PubMed: 19509336]
52. Ahuja U, Kjelgaard P, Schulz BL, Thöny-Meyer L, Hederstedt L. Haem-delivery proteins in cytochrome *c* maturation System II. *Mol Microbiol*. 2009; 73:1058–1071. DOI: 10.1111/j.1365-2958.2009.06833.x [PubMed: 19682263]
53. Sutherland MC, Rankin JA, Kranz RG. Heme Trafficking and Modifications during System I Cytochrome *c* Biogenesis: Insights from Heme Redox Potentials of Ccm Proteins. *Biochemistry*. 2016; 55:3150–3156. DOI: 10.1021/acs.biochem.6b00427 [PubMed: 27198710]
54. Feissner R, Xiang Y, Kranz RG. Chemiluminescent-based methods to detect subpicomole levels of *c*-type cytochromes. *Anal Biochem*. 2003; 315:90–94. [PubMed: 12672416]
55. Rasband, WS. ImageJ. US Natl. Inst. Health; Bethesda MD: 1997. Available at <http://imagej.nih.gov/ij/>
56. Berry EA, Trumpower BL. Simultaneous determination of hemes *a*, *b*, and *c* from pyridine hemochrome spectra. *Anal Biochem*. 1987; 161:1–15. [PubMed: 3578775]

57. Eddy SR. A new generation of homology search tools based on probabilistic inference. *Genome Inform Int Conf Genome Inform.* 2009; 23:205–211.
58. Söding J. Protein homology detection by HMM-HMM comparison. *Bioinforma.* 2005; 21:951–960. DOI: 10.1093/bioinformatics/bti125
59. Kamisetty H, Ovchinnikov S, Baker D. Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proc Natl Acad Sci.* 2013; 110:15674–15679. DOI: 10.1073/pnas.1314045110 [PubMed: 24009338]
60. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinforma.* 2000; 16:404–405.
61. Yang Y, Gao J, Wang J, Heffernan R, Hanson J, Paliwal K, Zhou Y. Sixty-five years of the long march in protein secondary structure prediction: the final stretch? *Brief Bioinform.* 2016; doi: 10.1093/bib/bbw129
62. Gront D, Kulp DW, Vernon RM, Strauss CEM, Baker D. Generalized fragment picking in Rosetta: design, protocols and applications. *PloS One.* 2011; 6:e23294.doi: 10.1371/journal.pone.0023294 [PubMed: 21887241]
63. Song Y, DiMaio F, Wang RYR, Kim D, Miles C, Brunette T, Thompson J, Baker D. High-resolution comparative modeling with RosettaCM. *Struct.* 1993; 21(2013):1735–1742. DOI: 10.1016/j.str.2013.08.005

Highlights

- Direct evidence for heme trafficking in putative heme transporters is lacking
- A cysteine-heme crosslinking approach was developed in a cyt c biogenesis pathway
- Heme was trapped in CcmC and CcmE, directly identifying heme binding domains
- Stereospecific heme positioning and presentation by the WWD domain is demonstrated
- Cysteine-heme crosslinking approach has potential to establish heme transporters

Figure 1. The prokaryotic cytochrome *c* biogenesis pathways and trafficking of heme

A. Model of the holocytochrome *c* pathways and proposed trafficking of heme (red), with the putative heme binding WWD domain labelled in CcmC and CcmF. Apocytochrome *c* (yellow) possess the CXXCH motif for attachment to heme 2- and 4-vinyl groups. For simplicity, the thioredoxin proteins that reduce the CXXCH thiols are not shown.

B. Sequence of the WWD domain in System I (CcmC, CcmF) and System II (CcsBA).

C. Structure of heme. The Fischer numbering system was used to designate the vinyl groups.

D. Mechanism of spontaneous crosslinking of a cysteine thiol to a vinyl group of heme. For simplicity, a single cysteine thiol is shown. Red arrows represent a 2-electron transfer. Figure is modified from [18]. Engineered cysteines described here are proposed to form the cysteine/heme crosslink by this mechanism.

Figure 2. Heme binding site and cysteine crosslinking in CcmE

A. Sequence of CcmE*, missing its single TMD. All residues substituted for cysteines or discussed in the text are numbered. Conserved residues (red) and semi-conserved residues (cyan).

B, C. Apo-CcmE* structures (PDB:1SR3) with color scheme as in A. B, C represent two orientations, and the H130 residue where heme 2-vinyl is normally attached is labelled. Other residues that are labelled were substituted for cysteines in the H130A background. E132C is labelled, the only residue that crosslinked to heme out of 16 substitutions. Distance between proposed heme vinyl binding pockets are shown.

D. Heme (yellow) is modeled into the proposed heme 2-vinyl and 4-vinyl pockets in the orientation shown in “C”.

E. Affinity purifications of indicated CcmCDE complexes analyzed by SDS-PAGE, and probed with the indicated reagents.

F–H. UV-vis spectral analysis of purified (black) and reduced (red) CcmCDE complexes. The reduced α/β region from 500–600 nm is magnified.

I. Reduced pyridine hemochrome spectra. CcmCDE(WT), black; CcmCDE(H130A), green; CcmCDE(H130A/E132C), orange.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. The CcmC WWD domain crosslinks to heme via distinct cysteine substitutions: structurally mapping the heme binding site in CcmC

A. Schematic of CcmC's proposed function. Six transmembrane domains (green), conserved histidines (orange stars), WWD domain (green loop) and heme (red) are shown.

B. The 17 amino acids of the WWD domain were substituted with cysteine in the CcmCDE(H130A) background and CcmCDE complexes purified using an N-terminal GST tag. An initial screen was done and the ratio of heme in CcmC to free heme in the variants was determined by SDS-PAGE and heme staining (as shown in D) with wildtype CcmC set to 1. Green asteriks(*) indicate variants with a ratio of >2X wildtype, indicating that they have crosslinked heme with CcmC and these variants were further studied.

C, D. Three stable variants selected are shown in (C) Coomassie total protein stain and (D) heme stain. Gels are representative of three independent protein preparations.

E, F. The function for each CcmC variant determined by maturation of cytochrome c4 and compared to wildtype CcmC.

Figure 4. Analysis of heme and CcmE in CcmCDE complexes that are crosslinked to heme via CcmC cysteine substitutions

A, B. UV-vis spectra of 50 μ g of purified variants with Soret absorbance (~ 410 nm) used to determine total heme levels (in triplicate), quantified in triplicate (B).

C, D. Co-purification of CcmE with the CcmC WWD cysteine variants was analyzed by α -CcmE immunoblot (C), quantified in triplicate (D).

E–G. UV-vis spectra of purified (black) and reduced spectra (red) of the indicated crosslinked complex.

H. Pyridine hemochrome spectra, α -peaks maxima are labeled. The blue-shifted α -peaks of 553 nm for the CcmC(W114C)DE(H130A) complex (and others) compared to H130A (at 556 nm) confirms a single crosslink.

Figure 5. Double-crosslinking heme to CcmC (WWD domain) and CcmE (His130): structurally mapping the vinyl groups and heme stereochemistry in the WWD domain

A. Schematic of possible heme binding orientations in the WWD domain. Heme (red) with 2- and 4-vinyl groups, solid line indicates known covalent bond between the heme 2-vinyl and CcmE(H130) (i.e. in the CcmE WT), dashed lines indicate potential covalent crosslinks to heme.

B–E. CcmC WWD cysteine variants in the CcmCDE(WT) background were purified and assayed for (B) stability, (C) crosslinking with heme, (D) α-CcmE immunoblot and (E) α-GST immunoblot. Relevant bands are labeled. In lanes 3, a 72 kDa complex that contains GST tagged CcmC(W114C), heme and CcmE is shown.

F–H. UV-vis spectra of purified (black) and reduced (red) complexes with magnified α/β region from 500–600 nm.

I. Pyridine heme spectra with α-peaks labeled. CcmCDE(WT), black; CcmCDE(H130A), green; CcmC(W114C)DE(WT) blue; CcmC(D126C)DE(WT) orange; CcmC(R128C)DE(WT) magenta.

Figure 6. Double-crosslinking heme to CcmC (WWD domain) and CcmE (H130A/E132C): Mapping the orientation of heme in the CcmCDE complex

A. Cartoon of a double crosslink that is expected if the stereochemistry and position of heme in CcmC and CcmE is validated.

B–E. CcmC WWD cysteine variants in the CDE(H130A/E132C) (lane 2) background xor CDE(H130A) (lanes 1, 3, 4) background were purified and assayed for (B) stability, (C) crosslinking with heme, (D) α-CcmE immunoblot and (E) α-GST immunoblot. Relevant bands are labeled. In lanes 2, a 72 kDa complex that contains GST tagged CcmC (W114C), heme and CcmE(H130/E132C) is shown.

Figure 7. Co-evolution of CcmC and CcmD

Molecular modeling of CcmCD complex. CcmC - green, with the WWD domain (magenta) highlighted; CcmD - cyan. Key residues are labeled.

A, B. CcmC, CcmD interaction as predicted by co-evolved residues. The top 5 co-evolved pairs (CcmC/CcmD) are as follows: red - V102/P31; purple - V109/M27; blue - M118/Y17; Grey - L59/T28; orange - L227/Q42.

Figure 8. Structures of heme in CcmC, and in the CcmCDE complex

Molecular modeling of CcmCDE complex. CcmC - green, with the WWD domain (magenta) highlighted; CcmD – cyan; CcmE - sand. Key residues are labeled and crosslinked residues are labeled with an asterisk (*).

A. Predicted structure of the CcmC WWD domain with heme. Residues (W114C, W123C) crosslinked to the 4-vinyl are indicated by an asterisk (*).

B. Magnification of CcmC WWD domain heme with His ligands.

C. Magnification of CcmC WWD domain heme with conserved and crosslinked tryptophan residues.

D. Predicted ternary structure of CcmCDE with heme. CcmC and heme are rotated 180 degrees from (C) to visualize interaction with CcmE 2-, 4- vinyl pockets

E. Magnification of heme binding domains of CcmC, CcmE.