Article:

Adams, N.B.P. orcid.org/0000-0003-3080-3448, Bisson, C., Brindley, A.A. et al. (4 more authors) (2020) The active site of magnesium chelatase. Nature Plants, 6 (12). pp. 1491-1502. ISSN 2055-026X

https://doi.org/10.1038/s41477-020-00806-9

This is a post-peer-review, pre-copyedit version of an article published in Nature Plants. The final authenticated version is available online at: http://dx.doi.org/10.1038/s41477-020-00806-9.

The active site of magnesium chelatase

Nathan B P Adams[1,*,+], **Claudine Bisson**[1,+,3], **Amanda A Brindley**[1], **David A Farmer**[1], **Paul A Davison**[1], **James D Reid**[2], **and C Neil Hunter**[1,*]

[1]Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, UK
[2]Department of Chemistry, The University of Sheffield, Sheffield, UK
[*]N.B.P.A. nathan.adams@sheffield.ac.uk / C.N.H c.n.hunter@sheffield.ac.uk
[+]these authors contributed equally to this work
[3]Present address: Centre for Ultrastructural Imaging, New Hunt's House, Guy's Campus, King's College London, London. SE1 1UL

ABSTRACT

The insertion of magnesium into protoporphyrin initiates the biosynthesis of chlorophyll, the pigment that underpins photosynthesis. This reaction, catalysed by the magnesium chelatase complex, couples ATP hydrolysis by a ChlID motor complex to chelation within the ChlH subunit. We probed the structure and catalytic function of ChlH using a combination of X-ray crystallography, computational modelling, mutagenesis and enzymology. Two linked domains of ChlH in an initially open conformation of ChlH bind protoporphyrin IX and rearrangement of several loops envelops this substrate forming an active site cavity. This induced fit brings an essential glutamate (E660), proposed to be the key catalytic residue for magnesium insertion, into proximity with the porphyrin. A buried solvent channel adjacent to E660 connects the exterior bulk solvent to the active site, forming a possible conduit for delivery of magnesium or abstraction of protons.

The central metal ions in the cyclic tetrapyrrole-derived cofactors, magnesium in chlorophyll, cobalt in cobalamin, iron in heme, and nickel in F_{430}, are collectively critical for most living systems. These tetrapyrroles underpin photosynthesis, vitamin B12 biosynthesis, respiration and methanogenesis[1]. Despite their apparent simplicity, the insertion of each metal ion into its cognate macrocyclic ring requires surprisingly complex and poorly understood enzymes. These metal ion chelatases can also play a regulatory role in directing and controlling flux down various branches of tetrapyrrole metabolism. A comparatively good structural and mechanistic understanding exists for the relatively simple Class II chelatases[2–6], but our knowledge of the complex, multisubunit, Class I chelatases, typified here by magnesium chelatase, is more limited.

Much of the mechanistic work on this class of chelatases has focused on the reasonably tractable magnesium chelatases (MgCH; E.C.6.6.1.1) from bacteriochlorophyll and chlorophyll producing organisms. MgCH initiates the biosynthetic pathways for these pigments by inserting Mg^{2+} into the protoporphyrin macrocycle (Fig. E1) . MgCHs require at least three subunits; chlorophyll producing organisms have ChlI (~ 35 kDa), ChlD (~ 75 kDa) and ChlH (~150 kDa), and the closely related proteins from bacteriochlorophyll producing organisms are BchI, BchD, and BchH[7,8]. In cyanobacteria and higher plants a fourth regulatory subunit, Gun4, is required for full protein activity[9–11].

The genes for MgCH were originally identified and recombinant protein expression systems were developed some time ago[7,12], and extensive kinetic characterization of the MgCH has identified the roles of the subunits[13–18]. The current mechanistic and structural data suggest a model for the MgCH mechanism where the two AAA^+ subunits form the ChlID complex[16,19]. This complex transiently interacts with the body region of the ChlH protein via the C-terminal integrin domain of ChlD[20], then hydrolyses ATP, which drives a conformational change in the ChlH-porphyrin complex, promoting insertion of the Mg^{2+} ion into the protoporphyrin IX (P_{IX}) ring[21]. This sequence of events accounts for the kinetic data, but a complete model requires structural information for all three proteins, in particular the porphyrin binding subunit ChlH, where insertion of Mg^{2+} takes place. SAXS and electron microscopy of negatively stained ChlH, provided low resolution views of the ChlH protein, both free and bound to porphyrin[22], and the recently published crystallographic structure of apo-ChlH from *Synechocystis* at 2.5 Å[23] represented a significant advance in structural studies of this subunit. Based on the shape of the protein, together these studies show that ChlH consists of three major domains, described as the head, neck and body of the protein[17,22]. From the crystal structure, the body is described as 'cage-like', with several sub-domains surrounding an internal cavity postulated to be the substrate binding site[23], consistent with the demonstration that a mutant of ChlH lacking the head and neck regions can still catalyse metal ion insertion[17].

Given the lack of detailed structural or mechanistic evidence for the location of the porphyrin binding site or key catalytic residues, we have used a series of approaches to characterise the putative active site of magnesium chelatase. Here we present a high-resolution (1.6 Å) crystal structure of a 45 kDa truncated construct of *Thermosynechococcus elongatus* ChlH, a 2.5

Å resolution structure of full-length WT *Synechocystis* ChlH and the structures of several catalytically altered mutants. We identified a number of residues likely to be involved in substrate binding and catalysis, which were tested by site directed mutagenesis, kinetic studies and crystallography. This combination of approaches has allowed us to make a detailed proposal for the magnesium insertion step in chlorophyll biosynthesis.

Results

An inactive, truncated form of ChlH from *Thermosynechococcus elongatus* (T_ChlH_926) (residues P926-E1326) was designed based on a structural study of a corresponding fragment of the cobalt chelatase CobN subunit (Prof. M Warren, University of Kent, personal communication). A seleno-methionine derivative was used to experimentally derive phases and build a preliminary model, which was subsequently used to determine a 1.6 Å resolution structure of the native protein (Table E1). T_ChlH_926 is largely alpha-helical, with a small seven-stranded beta-sheet at its core (Fig. 1 **a**) that underlies a prominent surface groove (Fig. 1 **b**). T_ChlH_926 has 78% sequence identity with the corresponding region of full-length ChlH from *Synechocystis* sp. PCC 6803 and comparison with a previously published structure of this protein (PDB:4ZHJ)[23] shows that the surface groove observed in the T_ChlH_926 structure represents a portion of the putative active site cavity. After refinement of the polypeptide structure, inspection of the difference density map revealed two peaks within the surface groove that could not be accounted for by the protein or by water molecules. One of the peaks corresponded to a 6-coordinate metal which, based on ligand distances and size of the peak in comparison to the surrounding protein, was modelled as a potassium ion (Fig. 1 **d**). The ligands surrounding this site include three water molecules, the sidechain of K200 and the mainchain amides of G111 and A174 (metal-ligand distances 2.91 Å, 3.05 Å, 3.32 Å, 3.53 Å, 3.15 Å and 3.22 Å, respectively). In the structure of a similarly truncated form of ChlH from *A. thaliana* (PDB: 5EWU)[24] (72% sequence identity to T_ChlH_926 and RMSD Cα: 0.53 Å) (Fig. 1 **c-d**) this same site has been modelled to be occupied by a magnesium ion. The second difference peak sits in a slight depression between the sidechains of W56, T58, S110 and V112, indicating the binding of a small molecule that either exogenously co-purified with the protein or that originated from the crystallisation conditions. Intriguingly, a molecule of benzoate occupies the same position in the *A. thaliana* ChlH structure, with the carboxyl group of the benzoate forming hydrogen bonds with the mainchain amide of G1034 and T1035, and sidechains of T1035 and S1087 (equivalent to G57, T58 and S110, respectively, in T_ChlH_926), indicating the location of a putative carboxyl binding site (Fig. 1 **d**). Despite the high-resolution of the T_ChlH_926 structure, it is not possible to identify the bound ligand, which may be partially disordered. Nevertheless, the conservation of both the architecture of this region and the surrounding residues (Fig. 1 **d**), suggests that this site may be important for substrate binding. We therefore hypothesised that this could be the binding site of one of the two pendant propionate groups of the ChlH substrate, P_{IX}, assigning it as site 1.

To get a better overview of the structure of the putative active site of ChlH and identify the second propionate binding site, we determined a 2.5 Å structure of full-length ChlH from *Synechocystis* sp. PCC 6803 (Syn_ChlH) (Table E1) by molecular replacement using a previously published 2.5 Å model of the enzyme (PDB:4ZHJ)[23]. Model building was aided by our high-resolution structure of T_ChlH_926, which represents 30% of the scattering matter in the crystal. The structure of Syn_ChlH is best described as comprising four domains (Fig. 2 **a**). The head (M1-P215) and neck (P232-D337) domains, connected by a long loop that is not visible in the structure, are packed against the body domain, which is formed of two sub-domains; Body I (A338-P942) and Body II (T950-V1330). The two halves of the body comprise roughly equally sized portions connected by a loop (A942-T950) (Fig. 2 **b**). Body II is the part of the full-length structure represented by the T_ChlH_926 truncate, whilst Body I forms a C-shape structure that embraces Body II (Supplementary Fig. S1 **a**), forming the compact cage-like structure described previously[23]. The active site cavity lies at the interface between the two body domains and is enclosed on all sides by protein, with a loop (Loop 1) (G466-L474) from Body I blocking access to the external solvent at the top of the cavity (Fig. 2 **c** and Fig. 3 **b**). Loop 1 and a number of surrounding loops, also from Body I (Loop 2: F618-K630 and Loop 3: Q575-D583), have varying degrees of order across the multiple copies of Syn_ChlH represented in our set of WT and mutant structures, suggesting that they are flexible.

There are two prominent electron density features (5 and 8 σ) sat deep within the active site cavity that could not be accounted for by the protein or by solvent structure, indicating the binding sites of unknown ligands (Fig. 2 **b-c**). The first corresponds to site 1 from the T_ChlH_926 structure, analogous to a site for hydrogenobyrinic *a*,*c*-diamide (HBAD) binding in a truncated CobN subunit (Prof. M Warren, University of Kent, personal communication). The second electron density feature (site 2), around 11 Å away, corresponds to the location of the bound potassium ion in the T_ChlH_926 structure. Site 2 is surrounded by a cluster of residues including S1103, K1129, D1177A and S1178A (Fig. 2 **c-d** and Fig. 3 **c**). Each electron density feature sits above the N-terminal end of a short helix and has a pair of exposed main chain amides directly beneath it (site 1; T987 and G986, site 2; V1041 and G1040), which are well positioned hydrogen bond donors to interact with the two propionate groups of the substrate (Fig. 2 **d**). The inside surface of the active site cavity around this region is largely positively charged, higher up it is neutral and at the top of the cavity next to Loop 1 it is more negatively charged, creating a charge gradient across the inside of the cavity (Supplementary Fig. S1 **b**). The resolution of the structure is insufficient to directly identify the

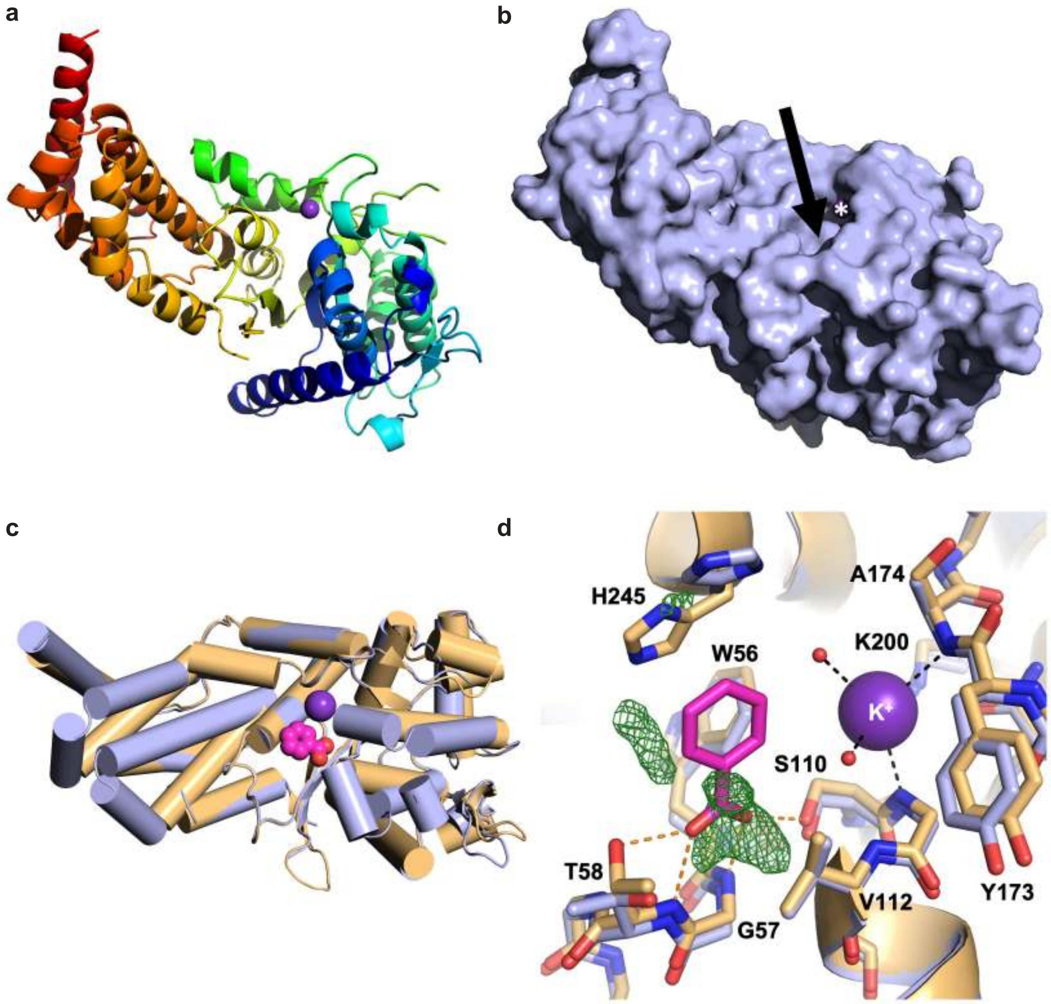

Figure 1. **a**, A cartoon representation of the *Thermosynechoccus elongatus* ChlH truncate, T_ChlH_926, coloured from the N-terminus (blue) to the C-terminus (red). A bound potassium ion is shown as a purple sphere. **b**, A surface representation (blue) of T_ChlH_926 in the same view as **a**. The black arrow indicates the location of the surface groove containing the potassium ion (white asterisk). **c**, A superposition of T_ChlH_926 (blue) with the structure of a similar truncate of ChlH from *A. thaliana* (beige), showing their conserved fold. The potassium ion bound to T_ChlH_926 and the benzoate molecule from the *A. thaliana* ChlH structure are shown as a purple sphere and as magenta spheres, respectively, indicating the conserved position of the surface groove. Helices are represented as cylinders. **d**, A detailed view of the binding groove of T_ChlH_926 (blue) superimposed on that of *A. thaliana* ChlH (beige) showing the conservation of amino acids in this region (numbering taken from the T_ChlH_926 structure) and the interactions that they make with the ligands in each structure. Also shown is a section of the Fo-Fc difference map from the T_ChlH_926 structure (green mesh, contoured at 3 σ) indicating the binding site of an unidentified ligand, the position of which corresponds to the benzoate in the *A. thaliana* ChlH structure, suggesting a conserved carboxyl binding site. We have termed this putative propionate binding site, site 1.

ligands, so to rule out an endogenous source we carried out gentle denaturation, refolding and HPLC of the enzyme prior to crystallisation. The features persisted (data not shown), leading us to believe that they must originate from the crystallisation conditions. Small carboxylic acids fit within the electron density features (butanoic or propanoic acid) but neither are present in the crystallisation, so we assume that these electron density features represent partially ordered molecules of citrate. Citrate inhibits magnesium chelatase activity ($K_{\mathrm{i,app}} \approx$ 80 mM Supplementary Fig. S2) and although its binding is relatively weak, at high concentration ($\approx$ 1 M citrate in the crystallisation experiment) it is feasible that Syn_ChlH-citrate complexes would form. We attempted to crystallise Syn_ChlH without citrate and with a number of different substrate analogues, both by soaks and co-crystallisation, but were unsuccessful. We believe that a "closed" form of Syn_ChlH crystallises, which is similar to that previously observed in a low-resolution ($\approx$ 30 Å) electron microscopy maps[22] and that perhaps the crystallisation is driven by citrate binding. The relative positions of sites 1 and 2 and the presence of a pair of main chain amides underlying each site (Fig. 2 **d**) suggests that these could be the binding sites for the two propionate groups of $\mathrm{P_{IX}}$.

Computational modelling of substrate binding by ChlH

In order to determine if sites 1 and 2 were accessible by substrate, we carried out modelling studies. It quickly became apparent from manual docking that with the two propionate groups docked in sites 1 and 2, the position of Loop 1 makes the active site cavity too small to accommodate $\mathrm{P_{IX}}$ (Fig. 3 **a-b**). However, we predict that Loop 1 is flexible (Supplementary Fig. S3), therefore its conformation in our structure may not be representative of its conformation in the presence of substrate. From the relative positions of Loop 1 and the docked substrate, we hypothesise that Loop 1 interacts directly with the substrate via its A- and B-ring substituents, which could facilitate closure of the active site by induced fit. Given this hypothesis, we prepared the structure for computational modelling (AutoDock Vina[25]) by removing residues T470-Y473 from Loop 1, as well as water molecules, and adding polar hydrogen atoms.

The top hit from the computational modelling (Fig. 3 **a-c**) places the C-ring propionate group of $\mathrm{P_{IX}}$ in site 1 and the D-ring propionate group into site 2. However, due to the symmetry of the molecule across the C- and D-rings, $\mathrm{P_{IX}}$ can be also be modelled the other way round with the D-ring propionate in site 1 and C-ring propionate in site 2. This is because the atomic coordinates for Loop 1 were removed from the protein structure prior to the modelling, therefore the asymmetrical substituent groups on the A- and B-rings have nothing to pack against and thus, the orientation cannot be defined by the model. Nevertheless, the modelling supports our visual analysis, confirming that we have two propionate binding sites in the base of the Syn_ChlH binding pocket. The hydrogen bonding network around each of the propionate groups is not clearly defined by the computational modelling, which has its limitations, but the position of the propionate groups suggests that they interact with the protein via the pair of main chain amides that underlie each site (site 1; T987 and G986, site 2; V1041 and G1040) as well as some of the surrounding side chains. Thus, we defined a set of residues around each of the putative binding sites to mutate (Site 1: S1039, V1041, T987 and D988, Site 2: S1103, K1129, D1177 and S1178, Table E2, Supplementary Table S1 and Fig. 3 **e-f**). All these residues are positioned on the Body II domain of Syn_ChlH and are conserved in ChlH from cyanobacteria, algae and higher plants, but there is less conservation with the related cobalt chelatase HBAD binding subunit, CobN (Fig. 3 **d**, Supplementary Fig. S5). Prior to any structural information, we also identified two other potential carboxyl binding residues (R667 and K991) that could be involved with substrate binding based on sequence conservation with CobN (Table E2). In order to determine if any of these residues were involved with substrate binding or catalysis we subjected the mutant enzymes to kinetic assays.

Activity and substrate binding assays in Syn_ChlH mutants

Interpreting the functional role of individual residues using steady-state kinetic measurements is challenging because of the complexity of the magnesium chelatase reaction mechanism. Three subunits (H, I and D) form the active complex and multiple enzymatic steps occur, so it is clearly inappropriate to assume that individual steady-state parameters reflect clearly defined reaction steps. They do however provide a clear overview of the effect of structural change. It is also possible to directly measure porphyrin binding affinity to ChlH[26] and in the analysis of our mutant enzymes we have determined binding constants (K_{d}) for $\mathrm{D_{IX}}$. In the analysis of the mutant enzymes we have therefore used a combination of porphyrin binding, measuring disruption of the active site, and steady-state kinetics to probe the overall ability of the mutant enzymes to catalyse the chelation reaction. In the analysis of steady-state kinetics, values for k_{cat} are calculated based on the concentration of ChlH.

Site 1 mutations

All four site 1 mutants (T987A, D988A, S1039A and V1041A) have impaired binding to substrate, with a higher $K_{\mathrm{d}}^{\mathrm{D_{IX}}}$ (2.3-6.3 μM) than WT ($\approx$ 0.65 μM) (Table E2 and Supplementary Fig. S4), consistent with our structural proposal. They also have lower turnover (k_{cat}) and lower catalytic efficiency ($k_{\mathrm{cat}}/K_{\mathrm{M}}$), compared to WT (Table E2). D988 sits underneath T987 and G986 (site 1 mainchain amides, see Fig. 2 **d**) and is involved in a network of water-mediated hydrogen bonds involving H1174. The D988 sidechain is fully hydrogen bonded in our structure (3 water molecules and S669-OH), therefore it may be structurally important or have a role in positioning waters in the active site. Interestingly, the T987A, S1039A and V1041A mutations

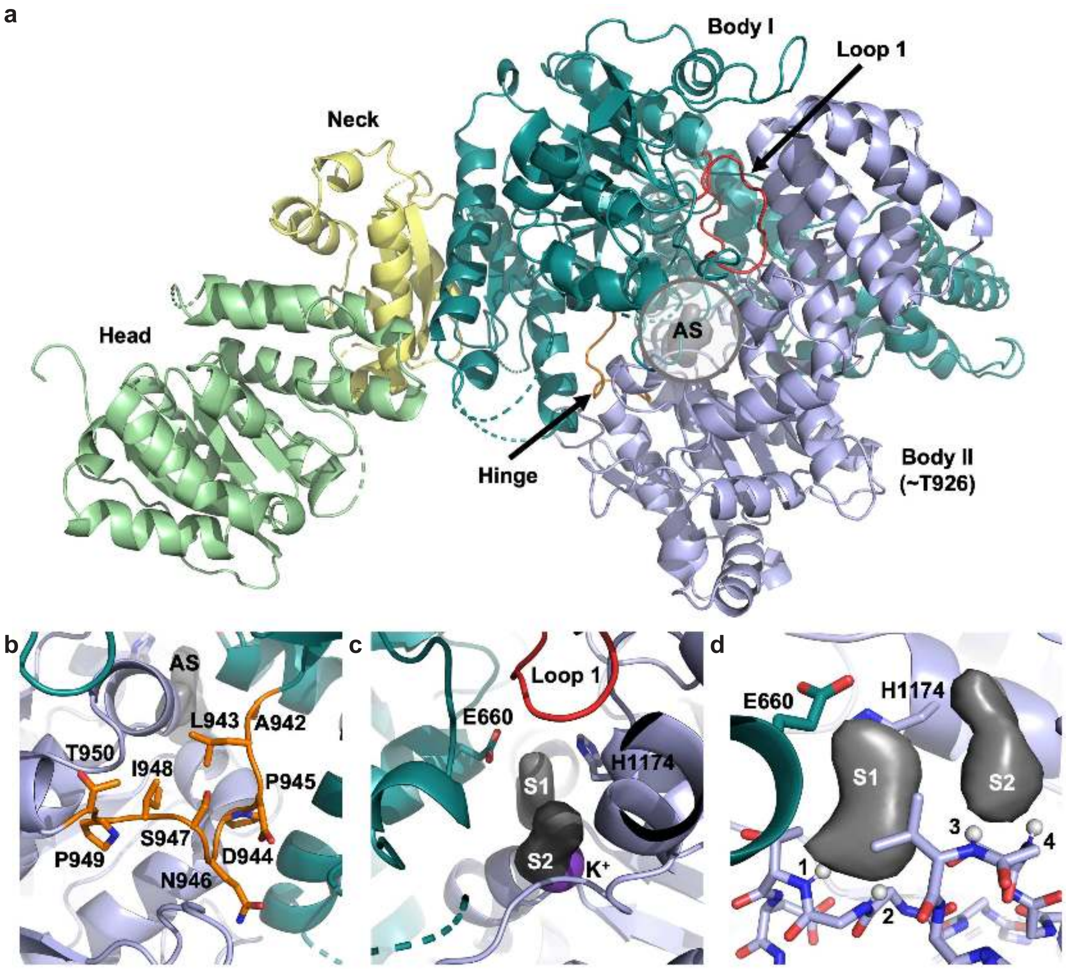

Figure 2. **a**, The three-dimensional fold of full-length Syn_ChlH represented as a cartoon. The head and neck domains are coloured green and yellow, respectively, with the body domain represented as two sub-domains, coloured according to the N-terminal portion (Body I; teal), which forms a C-shaped domain which cradles the N-terminal portion (Body II; pale blue). Body II is equivalent to the T_ChlH_926 structure (see Supplementary Fig. S1). The position of the active site (AS) is indicated by a grey circle; Loop 1 (red) and a potential hinge region (orange) are highlighted. **b**, The two sub-domains of the body are connected by a surface-exposed, hinge-like loop (orange) (A942-T950). **c**, Details of the active site cavity showing the difference density features (grey volumes) that indicate the binding of small molecules in sites 1 and 2. The potassium binding site from the T_ChlH_926 structure (purple sphere) that sits adjacent to S2 is overlaid and Loop 1 is shown in red. The sidechains of E660 and H1174 are also shown. **d**, Sites 1 and 2 are positioned at the N-terminal ends of two short helices, with a pair of mainchain amides sat directly underneath (small white spheres) (1=T987, 2=G986, 3=V1041, 4=G1040). These are well positioned hydrogen bond donors that could bind the two propionate groups of the substrate.

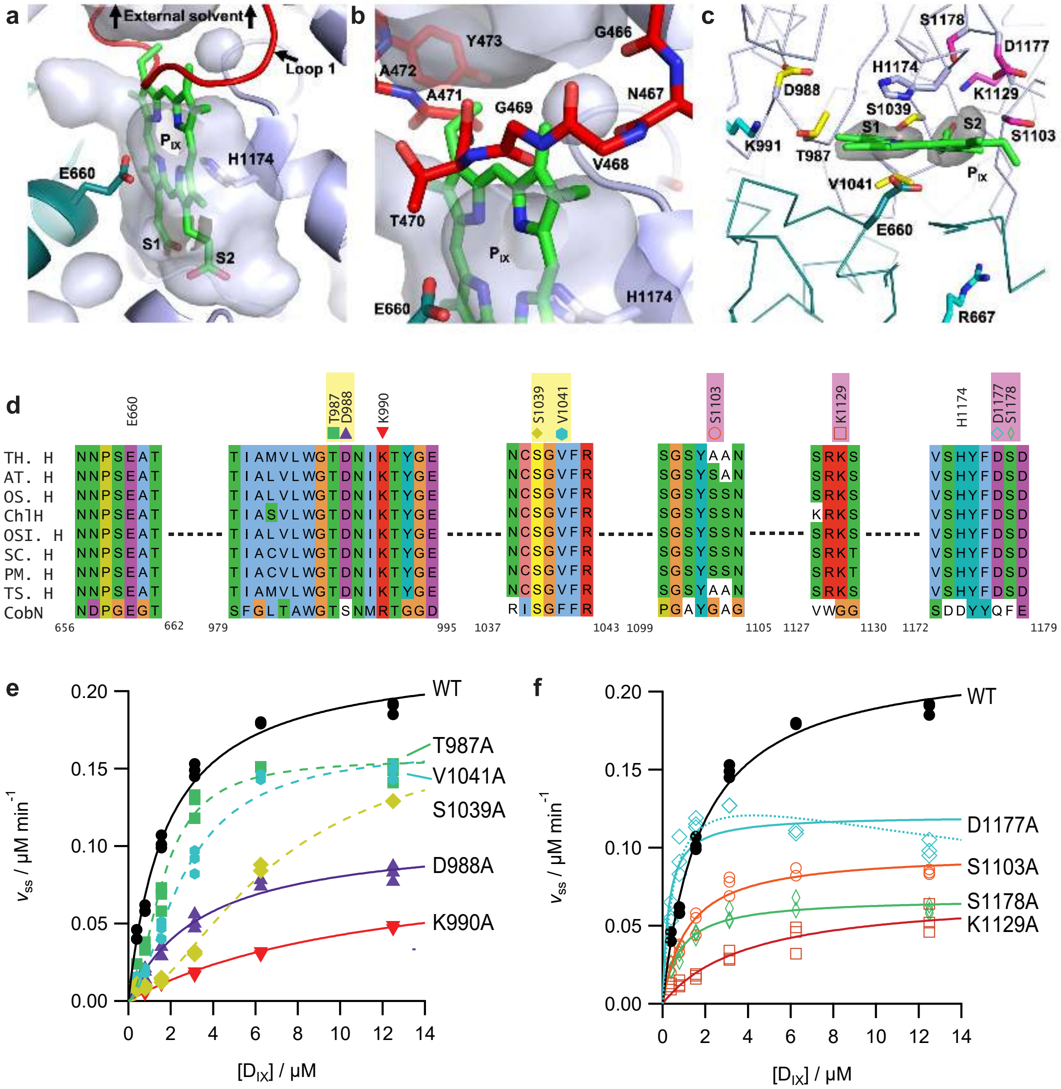

Figure 3. A model of P_{IX} binding in Syn_ChlH. **a**, P_{IX} (green sticks) sits within the central cavity of Syn_ChlH (partially transparent grey surface) with its two propionate groups docked in site 1 (S1) and 2 (S2). E660 sits on one side of the ring on the Body I domain (teal) and H1174 sits opposite it on the Body II domain (pale blue). The position of Loop 1 (red) is not compatible with the modelled position of the substrate, suggesting that its conformation would have to be different in a substrate-bound form of the enzyme. **b**, A more detailed view of Loop 1 (red) (shown as sticks and with residues labelled). Body domain colours are the same. **c**, A top-down view of the active site with P_{IX} (green) docked. Potential substrate binding residues around site 1 (yellow), site 2 (pink) and from sequence alignments with CobN (cyan) are shown as sticks on a ribbon backbone. E660 and H1174 are also shown. Sites 1 and 2 are represented by the difference density features observed in the full-length structure (grey partially transparent surface). Body domain colours are the same. **d**, Sequence alignment and secondary structure representations of the sequence highlighting the level of conservation amongst ChlH proteins and CobN for these core residues. Sequence alignment is coloured using the Clustal X colour scheme. TH. H, *Thermosynechococcus elongatus*; AT. H, *Arabidopsis thaliana*; OS. H, *Oryza sativa subsp. japonica*; ChlH, *Synechocystis sp. 6803*; SC. H, *Synechococcus sp. (strain WH8102)*; PM. H, *Prochlorococcus marinus (strain MIT 9301)*; TS. H, *Thermosynechococcus sp. NK55a*; CobN, *Brucella melitensis*. Numbering represents Syn_ChlH residue numbers. **e**, Steady state magnesium chelatase assays with respect to D_{IX} for site mutants. Assays contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT or mutant ChlH in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 5 mm $MgATP^{2-}$, pH 7.7. The lines are theoretical, with steady state rates described by fitting to the equations as listed in Table E2, along with fit coefficients. **f**, as for **e**, but for Site 2 mutants. The D1177A mutant (cyan) shows fits to both the Michaelis-Menten model (solid line) and substrate inhibition model (dashed line). In panels e and f, each data point is an independent experiment.

change the enzyme kinetics from a standard Michaelis-Menten rectangular hyperbola to a sigmoid, that can be described by the Hill equation with n of 1.7-1.9, which could suggest cooperativity. In our model, these three residues are predicted to hydrogen bond with the C-ring propionate of P_{IX}, either by mainchain or sidechain interactions. The positive cooperativity induced by these mutations are difficult to rationalise, but could explained by local structural changes around site 1 or by longer-range effects[27]. Overall, the subtle increase in K_d we observe for these mutants reflects the contribution of the WT sidechains to substrate binding in addition to the pair of mainchain amides that underlie site 1, agreeing with our model.

Site 2 mutations

We probed site 2 by making S1103A, K1129A, D1177A and S1178A mutations. As well as the mainchain amides from V1041 and G1040 that underlie site 2 (Fig. 2 **d**), the computational modelling suggests that the mainchain amide from S1103 could also act as a hydrogen bond donor to the D-ring propionate of P_{IX}. Mutating S1103 to alanine does perturb binding and turnover somewhat, but the effects are far more dramatic with the K1129A mutant, which has *ca.* 10% of the catalytic efficiency of WT, and an order of magnitude weaker binding to D_{IX}. In our structure, D1177 and S1178 sit in a constellation around K1129, with their sidechains pointing towards the molecule of docked P_{IX}. The space between the sidechains forms a small cavity adjacent to site 2. The N-zeta atom of the amino group of K1129 is ≈ 4.5 Å away from the nearest oxygen atom of the D-ring propionate group and a very subtle rearrangement could bring it within hydrogen bonding distance of the substrate. The big increase in K_d that we observe for the K1129A mutant is consistent with a direct interaction with the propionate group in site 2. For D1177 and S1178 a greater degree of movement would be required, so perhaps they facilitate water mediated interactions with the substrate or interact with the substrate in a different conformation of Syn_ChlH. Both the D1177A and S1178A mutants retain tight binding of D_{IX} (Table E2). Intriguingly, mutating D1177 to an alanine displays behaviour reminiscent of substrate inhibition, with a K_i value of 37 μM and a K_M value consistent with tight binding to D_{IX} during catalysis. This substrate inhibition behaviour has been previously observed with the *Themosynechoccus elongatus* enzyme[28]. The D1177A mutation may increase non-productive binding of substrate to ChlH leading to the observed substrate inhibition[29]. D1177A still has quite a high turnover number, and as such is actually more catalytically efficient than WT. The S1178A mutant also has a lower K_M and K_d (almost three times tighter binding to substrate than WT), but due to k_{cat} being significantly lower than WT, the overall effect is for a less catalytically efficient enzyme. Overall, the results of these mutational and kinetic analyses support the computational modelling and our hypothesis that this region is responsible for binding the D-ring substituents of P_{IX}.

Mutation of conserved positively charged residues

Prior to the determination of the full-length structure of Syn_ChlH, we used the structure of T_ChlH_926 along with sequence conservation to identify R667 and K991, conserved positively charged residues that could be near to the active site. Mutation of either residue to alanine in Syn_ChlH reduces the binding affinity of substrate, (K_d for D_{IX} 9.26 μM and 4.53 μM, respectively, compared to WT, $\approx$ 0.65 μM) (Table E2 and Supplementary Fig. S4), but whilst the K991A mutant still retains some enzyme activity, the R667A mutant is completely inactive; the same is true for R667K and R667E mutants. Analysis of the Syn_ChlH structure shows that R667 is surface exposed and forms an intramolecular hydrogen bond to Y1102 that bridges across the domain interface between the two halves of the body. From its position in the structure it is unlikely to interact with the substrate directly, but as mutation of this residue abolishes activity (Table E2) it is clearly critical, perhaps serving as a latch that stabilises the closed form of the enzyme. K991 is an important structural residue, with its sidechain amino group making three hydrogen bonds to a water molecule, sidechain of D1163 and the mainchain carbonyl on N656, respectively. It serves as a linchpin in the hydrogen bond network on one side of the active site cavity, and therefore plays a structural role. Both residues are important for enzyme activity in Syn_ChlH, but neither seems to be involved with substrate binding directly.

Residues critical for catalysis in the Syn_ChlH active site

Having shown that the mutational studies around sites 1 and 2 are consistent with our computational model of substrate binding, we sought to identify nearby residues that could be involved with catalysis. With the two propionate groups of the modelled substrate docked at the base of the active site, the porphyrin ring is sandwiched closely between the two walls of the cavity. On one side of the cavity, H1174 sits adjacent to the center of the porphyrin ring and on the opposite wall, projecting toward the other face of the porphyrin ring, is a strictly conserved glutamate residue (E660) (see sequence alignments in Fig. 3). This generates a Glu-His pair that form an axis (Glu-O to His-N = 6 Å) near to the predicted position of the magnesium binding site of the porphyrin ring (Fig. 2 **c** and Fig. 3 **a**).

An axial histidine would be well placed to either act as a base that abstracts protons from the porphyrin substrate and/or interact with the central Mg^{2+} ion of the product. Histidines have a catalytic role in a number of class 2 chelatases, including CbiK (an anaerobic cobalt chelatase from *Salmonella typhimurium*;[30] and *Bacillus subtilis*, human and *Saccharomyces cerevisiae* protoporphyrin IX ferrochelatases;[31–33]). H1174 is conserved across the entire magnesium chelatase family (Supplementary Fig. S8), but not in the analogous subunit in the cobalt chelatase, CobN, where it is replaced by an aspartate, perhaps reflecting the difference in metal ion specificity. In ferrochelatase, the equivalent axial histidine (H263, human protein residue numbering) is hypothesised to assist in proton abstraction from the porphyrin and mutation of this residue leads to severely impaired or inactive enzyme[34,35]. In Syn_ChlH mutation of H1174 to a valine showed no decrease in chelatase activity (Fig. E2 **a & b**, Table E3 and Supplementary Table S1), nor was the K_d for D_{IX} binding significantly altered (K_d = 0.77 μM compared with 0.65 μM for WT, Supplementary Fig. S6), demonstrating that it is not critical for binding or catalysis. Density functional theory calculations for Mg^{2+} insertion into porphyrin show a steep increase in the energetic challenge of deprotonation of the porphyrin ring by a water molecule as opposed to a histidine base[36]. We conclude, therefore, that H1174 is not catalytic, but it is in a convenient position to interact with the central magnesium. As histidine is a common ligand binding moiety in many chlorophyll binding proteins, perhaps it provides a stabilising interaction to anchor the Mg^{2+} ion until MgP_{IX} is released, but more evidence is required to confirm this.

The position of E660 on the other side of the porphyrin ring suggests that it could also interact with the central magnesium ion during catalysis. Accordingly, we made several mutations, including E660R, E660H, E660D and E660Q (Supplementary Fig. S8). Assays showed (Fig. E2 **c & d**, Supplementary Figs. S6 & S8, Table E3) that only by retaining a carboxyl group at position 660, via an aspartate residue, could any activity be retained (5 % compared to WT). We also determined a structure of E660D Syn_ChlH, showing that it is essentially the same as WT (RMSD Cα: 0.153 Å) (Table E1). The E660D mutant was able to bind D_{IX} with a K_d well within the range of other active mutants of ChlH, (E660D K_d = 3.05 μM, WT K_d = 0.63 μM), so the active site is still folded and able to bind D_{IX}, but the k_{cat} value for the E660D mutant is dramatically reduced compared to WT (Table E3). The E660D mutation produces a large increase in K_M with respect to D_{IX}, which is likely due to a decrease in catalytic proficiency of the enzyme, clearly demonstrated by the large decreases in the value for k_{cat}/K_M (or k_{cat} / $S_{0.5}$ for Mg^{2+}). This shows that the carboxyl group of E660 is essential for catalysis in Syn_ChlH and its location, adjacent to the center of the porphyrin ring, suggests that E660 plays a crucial role in inserting the Mg^{2+} ion into P_{IX}.

To further demonstrate that E660 sits adjacent to the substrate, we mutated it to a tryptophan (E660W) and carried out Förster Resonance Energy Transfer (FRET) assays, monitoring emission from D_{IX} following excitation of tryptophan at 280 nm[15]. The E660W protein still binds D_{IX} and as with the majority of the E660 mutants is inactive (K_d = 0.30 μM, Supplementary Fig. S7). The crystal structure of ChlH_E660W (Table E1) shows that the fold is very similar to WT (RMSD Cα: 0.30 Å). All of the putative substrate binding residues within the Body II domain of the enzyme are in conserved positions (Fig. 5) and the tryptophan is tolerated by packing the face of the indole ring against the wall of the active site, with some local rearrangement of adjacent loops (Loop 1-3) (see Fig. 5 **c-d**). The D_{IX} fluorescence emission from the E660W mutant was 47% greater than from the WT protein, while a control mutation, E263W, where the tryptophan is *ca.* 50 Å from the proposed porphyrin binding site, only showed a 7% increase in fluorescence (Fig. E3). Thus, we conclude that the substrate binds in close proximity to W660 providing direct evidence that the central cavity is the binding site of P_{IX}. Finally, Trp fluorescence quenching experiments have suggested that Syn_ChlH may bind substrate in a distorted non-planar conformation[26] and several crystal structures of unrelated proteins, including PDB:4UC1 and PDB:5DUO (both *Rhodobacter sphaeroides* translocator proteins) have non-planar porphyrins bound. Visual analysis of the active site cavity of Syn_ChlH shows that there is a lateral bend, such that a bowed substrate would bind with the convex face of the porphyrin ring pointing towards E660, presenting the four central nitrogen atoms of the porphyrin to the carboxyl side chain. This suggests that bending of the substrate within the Syn_ChlH cavity might facilitate metal insertion, although more work is required to confirm this.

Insights into magnesium ion binding in ChlH

Chelation of magnesium by P_{IX} requires a delivery mechanism to supply magnesium ions to the active site. This could either be achieved by direct uptake of magnesium by ChlH or an, as yet unknown, direct delivery mechanism from ChlD or ChlI. None of our Syn_ChlH structures have sufficient resolution to observe bound magnesium ions, a problem compounded by the complexity of the difference map features in the active site arising from binding of citrate. However, examination of the section of difference map adjacent to E660 reveals an intriguing chain of four buried electron density features sitting behind Loop 2 (F618-K630 ordered, disordered to P640) indicating a solvent-filled channel that leads from the surface of the protein to the active site. Although the surrounding residues are tightly packed against the buried solvent, a degree of local motion would make parts of this channel wide enough to accommodate a partially hydrated magnesium ion, notwithstanding the fact that this region of the enzyme may undergo substantial conformational change when in complex with ChlD and ChlI.

Dummy atoms were modelled into the centre of each electron density feature to aid interpretation, but these are not present in the deposited coordinates. The closest feature to E660 (feature 1) (Fig. 4 **a-b**) sits in a largely hydrophobic environment, with only the C-O of E660 and the next feature (feature 2) sat within hydrogen bonding distance. Feature 2 is surrounded by another three interacting groups, including the C-O of G622, -OE2 of E625 (both 2.7 Å away) and -OH of Y653, which is 2.2 Å away. This is too close for a water molecule and could be consistent with an O-Mg distance, notwithstanding the error in the map due to the limited resolution. The -OE1 of E652 points away from the solvent filled channel, hydrogen bonding to the -OH group of Y651. The -OE1 group of E625 also sits at the interface with bulk solvent, via a second empty channel that penetrates deep inside the body of Syn_ChlH, and is within hydrogen bonding distance to feature 3, along with the main chain amide of

H621. Feature 4 sits 5 Å away from feature 3, also at the interface with bulk solvent. Two groups are within hydrogen bonding distance of feature 4, the C-O of T457 and main chain amide of F618.

Y651, Y653 and E625 are strictly conserved across the ChlH family suggesting they are important (Fig. 4 **c**, Supplementary Fig. S5), thus to explore a potential role in metal binding, we made a series of mutations at the Y653 and E625 positions. Mutation of Y653 to a threonine still allows for substrate binding with a similar K_d to WT, but only reduces k_{cat} by ~40 %, suggesting that Y653 is not essential for supplying magnesium to the active site. To explore a possible role for E625 in metal binding we made E625D, E625H, E625K and E625Q mutants, all of which produced normally folded protein (Supplementary Fig. S9) and bound substrate with a K_d similar to WT, but with abolished chelatase activity. Due to this complete loss of activity it is difficult to be certain if E625 has a role in metal delivery, but from the conformation of the protein in our structure we can rule out a direct interaction with the substrate, since E625 is ~8 Å away from E660 (closest O to O distance) and points away from the active site. Given that the E625D mutant is also completely inactive, this suggests that there is also no tolerance for altering the position of the carboxyl group. We managed to obtain a crystal structure of the E625K mutant to 2.8 Å resolution (Table E1). Most of the structure is identical to the WT enzyme, including all of the substrate binding residues surrounding the propionate binding sites on the Body II domain and the position of the E660/H1174 pair. However, the K625 sidechain is pulled into the active site, altering the conformation of Loop 2 (F618-P640, disordered in the structure after residue 267), and obliterating the buried solvent channel. The conformational change of Loop 2 also destabilises the packing arrangement around Loop 1, the path of which is difficult to interpret from the electron density map. These conformational changes alter the shape of the active site cavity, creating more volume adjacent to E660. Interestingly, both the E625K and E660W structures show that a large degree of conformational change can be tolerated by the loops that overlie this buried channel (Loops 1-3), without causing much change in the positions of the substrate binding residues on the other side of the binding pocket. Consequently, they still bind substrate as well as WT.

Acidic residues adjacent to E660 would be likely candidates to be involved with vectoral delivery of magnesium to the active site, thus the position of E625 and its access to bulk solvent via a buried channel is consistent with a potential role in magnesium delivery. However, a route for protons to exit the active site cavity is also a requirement of the magnesium chelatase mechanism and a chain of buried solvent adjacent to E625 could also act as a proton wire, providing a path to bulk solvent for protons abstracted from the substrate prior to magnesium insertion.

Conclusion

The entire biosynthetic pathway for chlorophyll has been heterologously expressed in *E. coli*[37] and while we now know all the enzymes required to make this globally crucial pigment, we still lack a detailed mechanistic understanding of the individual reactions. The initiation and regulation of the chlorophyll pathway are important, and MgCH plays a pivotal role in both processes. Moreover, the chemistry of the MgCH catalyzed insertion of the Mg^{2+} ion into protoporphyrin is a particularly challenging and energetically demanding reaction. Synthetic routes to Mg-porphyrin products usually require anhydrous and oxygen-free environments, with high temperatures and long reaction times[38], in stark contrast to the constraints that operate in nature, which require insertion of magnesium into protoporphyrin in aqueous solvent. The enzyme must therefore sequester porphyrin away from solvent, generating a reaction vessel that can control the location of water molecules in the active site.

Magnesium chelatase is a large multi-subunit complex. The ChlH subunit binds to porphyrin in the absence of ChlID[22,26] and can associate to ChlID, either with or without porphyrin. The binding affinity between ChlD and ChlH is $\approx$ 330 nM[20], a far weaker interaction than between ChlD and ChlI ($K_d \approx$ 7 nM[19]). For a long time the ChlH-P_{IX} complex has been viewed as a substrate for ChlID[39].

Our work has used a combination of X-ray crystallography, *in silico* modelling of ligand binding and rational mutagenesis to probe the structure and catalytic function of ChlH. The binding site of the porphyrin substrate has been located, buried deep within the body domain of ChlH and located close to a catalytically essential glutamate (E660). We have shown from our structures that the bulk of ChlH is composed of two closely associated sub-domains (Body I and II) connected by a linking loop. Analysis of B-factors of the mainchain atoms show that Body II has on average much lower B-factors than Body I (Supplementary Fig. S3). Body II carries all of the substrate binding residues that surround sites 1 and 2 at the base of the active site cavity, whilst Body I possesses a number of flexible loops that have different conformations across our set of structures. Moreover, we can induce conformational changes in these loops via mutations (E625K and E660W) without decreasing substrate binding affinity, suggesting that they are not the main driver for recruiting substrate. We propose therefore that substrate can be loaded onto Body II with ChlH in a more open conformation, then the active site cavity can close around the substrate by rearrangement of the loops by induced-fit. Presumably the ChlI-D-$MgATP^{2-}$ complex is recruited to drive these motions by ATP hydrolysis. The insertion of magnesium into P_{IX} is thermodynamically unfavourable[14], but our previous work shows a considerable mismatch between the ΔG of Mg^{2+} insertion and the ATP stoichiometry of the reaction[14,19,40]. We previously suggested that this mismatch could be due to a probabilistic mechanism, where not every hydrolysis of $MgATP^{2-}$ brings about successful Mg^{2+} insertion. ChlH could undergo several conformational changes to facilitate catalysis and, while

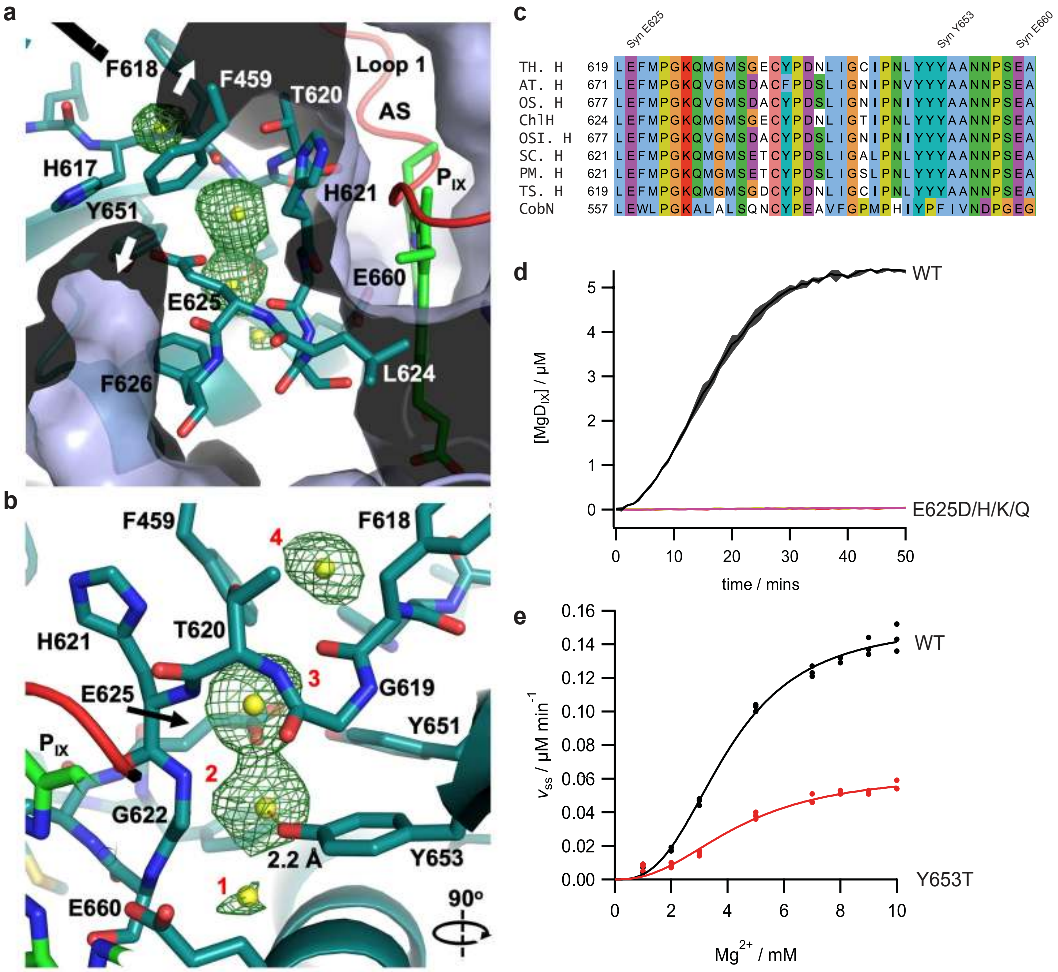

Figure 4. A chain of buried solvent features are present adjacent to E660 within the Syn_ChlH structure. **a**, The four buried positive difference features are shown as a green mesh (Fo-Fc map contoured at 3 σ) with a dummy atom (yellow sphere) modelled at the centre of each feature. Surrounding residues are drawn as sticks and labelled, the interior walls of the active site cavity (AS) are drawn as a partially transparent pale blue surface and the relative positions of the modelled P_{IX} (green) and Loop 1 (red) are shown for context. White arrows indicate two paths to bulk solvent, via a channel adjacent to E625 and on the surface of the protein, above F459 . **b**, The view in **A** is rotated 90° and much of the active site cavity is stripped away, showing a close approach (2.2 Å) between the Y653-OH and one of the buried electron density features, suggesting it might not be a water molecule. The features are labelled 1-4 in red text, corresponding to the description in the main text. **c**, Sequence alignment of ChlH and CobN. TH. H, *Thermosynecoccus elongatus*; AT. H, *Arabidopsis thaliana*; OS. H, *Oryza sativa subsp. japonica*; ChlH, *Synechocystis sp. 6803*; SC. H, *Synechococcus sp. (strain WH8102)*; PM. H, *Prochlorococcus marinus (strain MIT 9301)*; TS. H, *Thermosynechococcus sp. NK55a*; CobN, *Brucella melitensis*. **d**, Time courses of MgCH assays with WT and E625; assays were performed in triplicate and are displayed as the mean with stand deviations shown in shading. **e**, Mg^{2+} dependence of the steady state rate of Mg^{2+} chelation for WT and Y653T ChlH, lines are theoretical and described by Equation 6, with parameters: WT, $k_{cat}^{Mg} = 0.38 \pm 0.01$ μM min^{-1}, $S_{0.5}^{Mg} = 3.99 \pm 0.12$, $n = 2.8 \pm 0.2$ μM; Y653T, $k_{cat}^{Mg} = 0.16 \pm 0.01$ μM min^{-1}, $S_{0.5}^{Mg} = 4.41 \pm 0.38$, $n = 2.4 \pm 0.3$ μM. In panel **e** each data point is an independent experiment. Assays in panels **d** and **e** contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT or mutant ChlH in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 8 μM D_{IX}, 5 mm $MgATP^{2-}$, pH 7.7, unless otherwise stated.

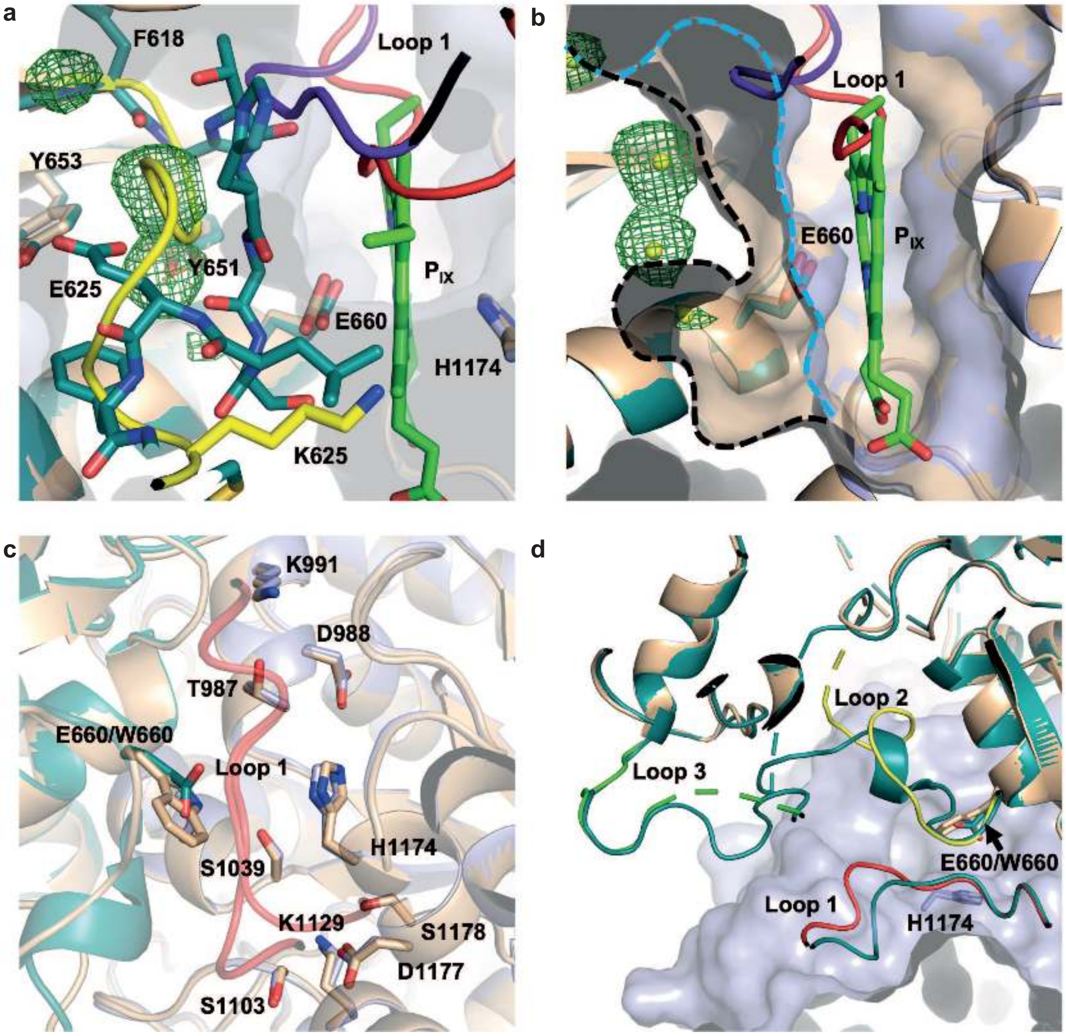

Figure 5. The E625K and E660W mutations alter the conformation of loops around the active site. **a**, The WT structure of Syn_ChlH is shown in teal (Body I), with a grey surface indicating the internal surface surrounding the modelled substrate (green) and Loop 1 shown in red. The structure of the E625K mutant (beige) is superimposed, with the loop bearing K625 (Loop 2; F618-P640 - becomes disordered in the E625K structure from residue 627) highlighted in yellow. The E625K mutation dramatically alters the conformation of Loop 2, which obliterates a buried solvent channel that is positioned behind E660 (indicated by the green mesh and small yellow spheres) (see Fig. 4 and corresponding text for a detailed description). Although much of the rest of the structure remains unchanged, including the position of E660 and H1174, the movement of Loop 2 causes the position of Loop 1 (purple) to alter. **b**, A comparison of the inside surface of the active site between the WT and E625K structures. The conformational change of Loop 2 alters the shape of the active site cavity in the E625K mutant, creating more volume adjacent to a buried solvent channel. This presumably alters the molecular environment around E660 (surfaces were generated with Loop 1 coordinates omitted); WT = blue and E625K = beige, cavity edge is indicated by a dashed line, in blue and black, respectively. **c**, A view of the active site cavity in the E660W mutant structure (beige), showing the position of the tryptophan sidechain compared to that of E660 in the WT structure (coloured by domain; Body I teal and Body II pale blue). All the putative substrate binding residues on Body II have conserved positions in the two structures. For context, the position of Loop 1 in each structure is also shown (red). **d**, A broader overview of the region surrounding W660, showing the conformational change in Loop 1 (red), Loop 2 (yellow) and Loop 3 (green) (Q575-D583), compared to their conformations in the WT structure (teal). All these loops are on Body I of Syn_ChlH. Body II is shown as a partially cut-away surface and the E660/H1174 pair is shown as sticks.

they may be thermodynamically freely accessible, $MgATP^{2-}$ hydrolysis could be required to accelerate progression through these conformations, accounting for at least some of the large ATP stoichiometry (15 - 40 molecules of ATP hydrolysed per MgP_{IX} formed).

The related class II enzyme protoporphyrin IX ferrochelatase (PpfC) might provide useful clues to the potential mechanism for magnesium insertion. Ferrochelatase is the most extensively studied member of the chelatase family and uses an identical porphyrin substrate, but inserts a ferric ion into P_{IX}. The large number of crystal structures of PpfC determined, including some with either native or non-native porphyrins bound, coupled with extensive kinetic analyses, has led to a detailed mechanistic proposal for the enzyme. In PpfC, a distal His abstracts protons from P_{IX} towards a patch of negatively charged residues, while a Met residue on the opposite face has been proposed to deliver Fe(II) ions to the substrate[35,41], so that enzyme residues on both faces of the porphyrin are involved in the reaction mechanism. Our kinetic analysis of the structurally similar H1174 in Syn_ChlH indicates that this residue is not catalytically essential and there is also a lack of negatively charged residues in the vicinity of the H1174. With the H1174 side of the active site not participating in the reaction, we propose a single sided mechanism for magnesium chelation, via E660, where both protein abstraction and magnesium ion insertion occur in a sequential manner on the same face of the porphyrin (Fig. 6) with the free energy from $MgATP^{2-}$ hydrolysis used to drive the conformational changes that to allow the reaction to proceed.

Methods and Materials

Cloning methods

The truncated version of *Thermosynechoccus elongatus* ChlH was engineered by introducing a new NdeI site into the DNA at Asp 925 using QuikChange mutagenesis with the primers detailed in Table S2. QuikChange PCR was carried out according to the manufacturer's protocol (Agilent, Cheshire, UK), using pET21aThermoH[22] as the template. The PCR product was digested with NdeI, gel purified and ligated with T4 DNA ligase (Promega). The resultant pET21aThermoH925 construct was verified by sequencing (GATC).

Site directed mutagenesis

Mutants of ChlH were generated using primers listed in Table S3 using the QuikChange method according to the manufacturer's protocol.

Protein production and purification

All full-length magnesium chelatase subunits and the T_ChlH_926 truncation were purified essentially as described previously[28]. Protein purity was estimated to be >95% as assessed by SDS-page. For gentle denaturation and refolding of the protein to attempt to remove co-purified ligands from the active site, 4 ml of ChlH was dialysed against 500 ml of 25 mM Tris, 2 M urea, pH 7.45, overnight at 4 °C. On day two, the protein was then refolded by dialysis with 20 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.45 and then finally overnight a final dialysis against the same buffer but with 2 mM DTT overnight, all at 4 °C. On day 3 the dialysis sample was diluted into 25 mL of buffer with 2 mM DTT freshly added, before concentrating to 250 μM for crystallisation trials.

Crystallisation and structure determination

The seleno-methionine incorporated construct representing the N-terminal third of *Thermosynechoccus elongatus* ChlH (SeMet_T_926) (residues 926 - 1331) (10 mg ml^{-1}) was crystallised by sitting drop vapour diffusion (1 μl drop, 1:2 ratio, 1 ml reservoir) in 20-22 % PEG 6000, 100 mM Tris, 200 mM $MgCl_2$, pH 8 (290 K). Cubic crystals formed after several days, which were harvested, cryoprotected in their mother liquor with an additional 25 % ethylene glycol, and flash-cooled by plunging into liquid nitrogen. Data were collected on a single crystal at the selenium edge (λ = 0.97901 Å) on beamline i24 at the Diamond Light Source. The crystal belonged to the spacegroup $P2_1$ with cell dimensions of a=76.48 Å, b=140.6 Å, c=78.59 Å and β=108.82 °. Data were processed with Xia2[42] using a CC-half resolution cut-off of 1.84 Å. Shelx[43] was used to locate ~30 selenium sites that had an occupancy of >0.5 and to build a poly-alanine starting model. Native crystals grew in the same conditions as the SeMet derivative in the presence of 10 mM bipyridine and a higher-resolution (1.64 Å) structure was determined by molecular replacement using the preliminary structure as a starting model. The structure was subjected to multiple rounds of manual and automated model building in COOT[44] and Buccaneer[45], respectively. The model was refined with Refmac5[46] and validated using Molprobity[47] (See Table E1 for data collection and refinement statistics).

The full-length construct of WT ChlH from *Synechocystis sp.* (ChlH) and a number of mutants (E600D, E660W and E625K) were crystallised by sitting-drop vapor diffusion (2 μl drop, 1:1 ratio, 1 ml reservoir) in 0.8-1.2 M sodium citrate pH 7.0 and 2 mM dithiothreitol (290 K). Large triangular rod-shaped crystals grew within several days, which were cryoprotected in Paratone N oil (Hampton Research) and flash-cooled with liquid nitrogen. Data from WT protein crystals were collected to 2.54 Å resolution and were processed with Xia2[42]. The crystals belonged to the spacegroup H3 with cell dimensions of

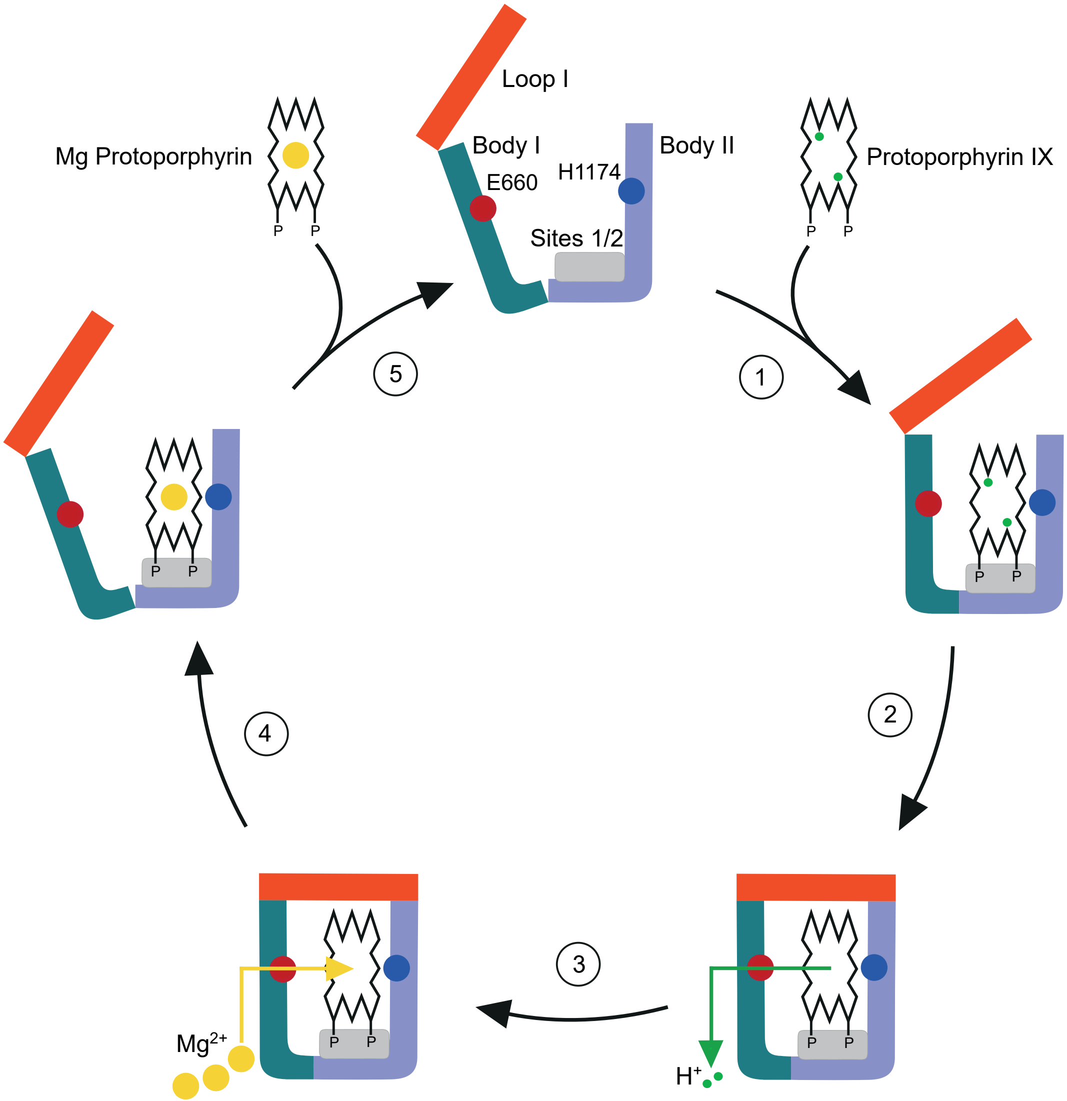

Figure 6. A proposed model for one-face sequential proton abstraction and magnesium ion insertion into P_{IX} by ChlH. From the top, the apo-ChlH active site is represented as a schematic cartoon, with the flexible Body I domain represented in teal and Body II in pale blue. The propionate binding region (sites 1 and 2) is in grey. Moving clockwise, (1) P_{IX} binds to ChlH by induced fit, with flexible loops on Body I packing around the substrate. Propionates on rings C and D bind to sites 1 and 2, and are denoted by the letters P. (2) Loop 1 (red) closes the active site cavity and protons are abstracted from pyrrole nitrogens via Body I (green arrow). (3) ChlH is remodelled by ChlD (not shown in diagram) and hydrolysis of $MgATP^{2-}$ drives Mg^{2+} delivery (yellow arrow) inserted into P_{IX} via a solvent-filled channel that leads from the surface of the protein to the active site, also via Body I. (4) The active site reverts to an open state, enabling (5) release of the MgP_{IX} product.

a=b=320 Å, c=105 Å, $\alpha=\beta=90$ °, $\gamma=120$ °. An initial model containing two chains in the asymmetric units was determined directly by refinement in Refmac5[46] (CCP4i[48]) using PDB:4ZHJ[23] as a starting model. As chain A was much better defined in the map than chain B, all of the modelling and structural interpretation was carried out using chain A. All the crystals of the mutant enzymes belonged to the same spacegroup and cell dimensions as the WT protein, so these structures were also determined directly by refinement. The resolutions of the mutant structures were somewhat lower than the WT ($\approx$ 2.8 Å, but in all cases difference density indicated the position of the mutated residue. Refinement was carried out in Refmac5 imposing secondary structure restraints generated by ProSMART[49] and latterly, using jelly-body refinement[50]. Model building was carried out in COOT[44], which was also used for model validation, along with Molprobity[47] (See Table E1 for data collection and refinement statistics). Images were made with PyMOL (V 1.2r1, Schrödinger, LLC). Structures have been deposited in the Protein Data Bank under accession codes: 6YS9, 6YSG, 6YT0, 6YTN and 6YTJ.

Ligand docking

Protoporphyrin IX (P_{IX}) docking into full-length *Synechocystis* ChlH was performed using AutoDock Vina[25]. An idealised structure of P_{IX} was downloaded from the PDB (PDB ligand code: PP9, extracted from PDB:2HRE) and hydrogens added using Open Babel (V.2.3.2,[51]). A general protocol from[52] and the AutoDock Vina website was adapted for this experiment. The full-length structure of ChlH was prepared for docking using AutoDockTools (MGLTools 1.5.6) where a PDBQT file was generated, and a 20 $Å^3$ grid for docking was located in the cavity surrounding site 1 (centered around residue E660, at position -29.2, 37.2, -33.6 for x, y and z for our structure).

Magnesium chelatase assays

Assays were performed as described previously[14]. Briefly assays were performed in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, I = 0.1 (KCl) at 34 °C with substrate concentrations of 15 mM $MgCl_2$, 5 mM ATP, 8 μM D_{IX}, unless otherwise stated in figure legends. Assays were initiated by the addition of enzyme to give final concentrations of 0.1 μM ChlD, 0.2 μM ChlI and 0.4 μM ChlH. Steady state assays of magnesium chelatase followed formation of the product, MgD_{IX}, over a period of one to two hours. Reaction progress was monitored using a Omega FluoStar microplate reader (BMG LabTech, Aylesbury, UK), with excitation at 420 ± 5 nm and emission detected at 580 ± 5 nm. Steady state rates (v_{ss}) were calculated using software supplied by the manufacturer (MARS version 2.41). Non-linear regression analysis was performed using Igor Pro (v. 8.04).

Citrate inhibition assay

Using a 1 M stock of sodium citrate (titrated to pH 7.9) standard magnesium chelatase assays were spiked within an increasing concentration of citrate (1.6 - 200 mM). As citrate is a known chelator of Mg^{2+}[53], the amount of $MgCl_2$ added to each reaction was equal to the concentration of citrate plus 15 mM (e.g. at 200 mM citrate the concentration of $MgCl_2$ was 215 mM). The apparent inhibition constant was calculated using Equation 7.

Porphyrin binding studies

Experiments were performed using a FluoroMax 3 spectrofluorometer (Jobin Yvon). Tryptophan quenching assays and calculations of K_d values for the binding of D_{IX} to ChlH were performed essentially as described previously[26], but due to the high concentrations of porphyrin required for weak binding mutants, a linear component was added to the binding equation[54] to take into account the inner filter effect of porphyrin absorbing light at high concentrations.

$$F_{obs} = F_0 + F_{min} \cdot \frac{[\mathrm{L}]_\mathrm{T} + [\mathrm{E}]_\mathrm{T} + K_\mathrm{d} - \sqrt{([\mathrm{L}]_\mathrm{T} + [\mathrm{E}]_\mathrm{T} + K_\mathrm{d})^2 - 4[\mathrm{L}]_\mathrm{T}[\mathrm{E}]_\mathrm{T}}}{2[\mathrm{E}]_\mathrm{T}} + M[\mathrm{L}]_\mathrm{T} \tag{1}$$

Enzyme in excess porphyrin binding studies

To determine the K_d of D_{IX} to ChlH E660W an enzyme in excess approach was used. Experiments were performed using a FluoroMax 3 spectrofluorometer (Jobin Yvon). 0.05 μM D_{IX} was mixed with increasing concentrations of ChlH WT or E660W protein in 50 mM MOPS/KOH, 0.3 M glycerol, pH 7.7 at 34 °C. D_{IX} was excited at 395 ± 5 nm and emission spectra recorded between 550 and 750 nm (5 nm bandpass). The intensity of emission at 611 and 672 nm was plotted, and both wavelength sets were simultaneously globally fit using equation 2 using the Global Fit package within Igor Pro.

$$F_{obs} = F_0 + F_{min} \cdot \frac{[\mathrm{L}]_\mathrm{T} + [\mathrm{E}]_\mathrm{T} + K_\mathrm{d} - \sqrt{([\mathrm{L}]_\mathrm{T} + [\mathrm{E}]_\mathrm{T} + K_\mathrm{d})^2 - 4[\mathrm{L}]_\mathrm{T}[\mathrm{E}]_\mathrm{T}}}{2[\mathrm{E}]_\mathrm{T}} \tag{2}$$

CD spectroscopy

Protein (0.1 mg ml^{-1}) was buffer exchanged into 5 mM sodium phosphate buffer, pH 7.4 and spectra were recorded in a cuvette with a 0.1 cm path length at 25 °C using a JASCO-810 spectrometer (JASCO, UK). Spectra were recorded continuously from 250 to 190 nm (50 nm s^{-1}, 1 nm increments, 4 s response, 6 accumulations) and background subtracted before calculation of mean residue elipticity (MRW) using Equation 3.

$$[\theta]_{\mathrm{MRW}} = \frac{MRW \times \theta}{10 \times c \times d} \tag{3}$$

Where c is the concentration in mg ml^{-1}, d is the pathlength in cm and MRW is the mean residue weight calculated using Equation 4.

$$\mathrm{MRW} = \frac{M_r}{(N-1)} \tag{4}$$

Fluorometric detection of enzyme-porphyrin complex

Experiments were performed using a FluoroMax 3 spectrofluorometer. Fluorescence emission spectra were recorded from 520 - 750 nm (10 nm slit width) with excitation at 280 nm (20 nm slit width). Emission was observed through a 500 nm long pass filter to exclude internally reflected incident light. 1 μM ChlH protein was incubated with 100 nM $\mathrm{D_{IX}}$ for 10 minutes at 34 °C in 50 mM MOPS/KOH, 0.3 M glycerol, pH 7.7, prior to spectra being recorded. Experiments were performed in triplicate.

Data analysis

Non-linear regression was performed using Igor Pro 8.04 (Wavemetrics). Kinetic curves were described by one of the Michaelis-Menten (Equation 5), Hill (Equation 6) or substrate inhibition (Equation 7) functions.

$$v_{ss} = \frac{V_{\max}[\mathrm{S}]}{K_{\mathrm{M}} + [\mathrm{S}]} \tag{5}$$

$$v_{ss} = \frac{V_{\max}[\mathrm{S}]^n}{s_{0.5}^n + [\mathrm{S}]^n} \tag{6}$$

$$v_{ss} = \frac{V_{\max}}{1 + \frac{K_{\mathrm{M}}}{[\mathrm{S}]} + \frac{[\mathrm{S}]}{K_{\mathrm{i}}}} \tag{7}$$

Where $v_{\max}$ is maximal velocity, K_{M} is the Michaelis constant, K_{i} is the inihibition constant, [S] is substrate concentration, n is the Hill coefficient. k_{cat} was calculated based on the concentration of ChlH within assays.

The apparent inhibition constant ($K_{\mathrm{i,app}}$) for citrate inhibition was calculated using Equation 8, an empirical inhibition function[55].

$$v_{ss} = \frac{v_{\mathrm{control}} - v_{\mathrm{bkg}}}{1 + \frac{[I]_0}{K_{\mathrm{i,app}}}} + v_{\mathrm{bkg}} \tag{8}$$

Where v_{control} is the steady state rate in the absence of inhibitor, v_{bkg} is the steady state rate at infinite concentration of inhibitor, which is held at 0 in the fitting and n is the slope factor or Hill coefficient.

Acknowledgments

The authors would like to thank Diamond Light Source for beamtime (proposals mx8987 and 12788), and the staff of beamlines i24, i02, i03 and i04 for technical support. The authors thank Dimitris Ladakis, Richard W. Pickersgill, David G. Brown and Martin J. Warren for the provision of structural data on the truncation of the CobN protein. N.B.P.A., C.B., A.A.B., P.A.D., J.D.R. and C.N.H. gratefully acknowledge financial support from the Biotechnology and Biological Sciences Research Council (BBSRC U.K.)[award number BB/M000265/1]. C.B. was also supported by core funding from The Department of Molecular Biology and Biotechnology, University of Sheffield. C.N.H was also supported by Advanced Award 338895 from the European Research Council. D.A.F. was supported by a BBSRC White Rose Doctoral Training Program [award number BB/M011151/1].

Author contributions statement

N.B.P.A. and D.F. carried out kinetics and biophysics experiments. A.A.B and P.A.D generated mutants and with N.B.P.A and D.F. prepared proteins. A.A.B, N.B.P.A and C.B. setup crystallisation experiments. C.B. determined, built and analysed structures. C.N.H provided laboratory space and materials to carry out the work. N.B.P.A, C.B. and C.N.H wrote the manuscript with contributions from J.D.R. The manuscript was approved by all authors.

Additional information

Data availability PDB accession codes: 6YS9, 6YSG, 6YT0, 6YTN and 6YTJ; raw kinetic and biophysical data available on request. **Competing interests** The authors declare that they have no competing interests.

References

1. Bryant, D. A., Hunter, C. N. & Warren, M. J. Biosynthesis of the modified tetrapyrroles–the pigments of life. *J. Biol. Chem.* DOI: 10.1074/jbc.REV120.006194 (2020).
2. Medlock, A., Swartz, L., Dailey, T. A., Dailey, H. A. & Lanzilotta, W. N. Substrate interactions with human ferrochelatase. *Proc. Natl. Acad. Sci. United States Am.* **104**, 1789–1793, DOI: 10.1073/pnas.0606144104 (2007).
3. Medlock, A. E., Carter, M., Dailey, T. A., Dailey, H. A. & Lanzilotta, W. N. Product release rather than chelation determines metal specificity for ferrochelatase. *J. molecular biology* **393**, 308–19, DOI: 10.1016/j.jmb.2009.08.042 (2009).
4. Hoggins, M., Dailey, H. A., Hunter, C. N. & Reid, J. D. Direct measurement of metal ion chelation in the active site of human ferrochelatase. *Biochemistry* **46**, 8121–8127, DOI: 10.1021/bi602418e (2007).
5. Hofbauer, S., Helm, J., Obinger, C., Djinović-Carugo, K. & Furtmüller, P. G. Crystal structures and calorimetry reveal catalytically relevant binding mode of coproporphyrin and coproheme in coproporphyrin ferrochelatase. *The FEBS J.* **n/a**, DOI: 10.1111/febs.15164 (2020).
6. Gillam, M. E., Hunter, G. A. & Ferreira, G. C. Ferrochelatase π-helix: Implications from examining the role of the conserved π-helix glutamates in porphyrin metalation and product release. *Arch. Biochem. Biophys.* **644**, 37 – 46, DOI: 10.1016/j.abb.2018.02.015 (2018).
7. Gibson, L. C., Willows, R. D., Kannangara, C. G., von Wettstein, D. & Hunter, C. N. Magnesium-protoporphyrin chelatase of rhodobacter sphaeroides: reconstitution of activity by combining the products of the bchh, -i, and -d genes expressed in escherichia coli. *Proc. Natl. Acad. Sci.* **92**, 1941–1944, DOI: 10.1073/pnas.92.6.1941 (1995).
8. Gibson, L. C., Jensen, P. E. & Hunter, C. N. Magnesium chelatase from Rhodobacter sphaeroides: initial characterization of the enzyme using purified subunits and evidence for a BchI–BchD complex. *Biochem. J.* **337**, 243, DOI: 10.1042/0264-6021:3370243 (1999).
9. Davison, P. A. *et al.* Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis. *Biochemistry* **44**, 7603–12, DOI: 10.1021/bi050240x (2005).
10. Mochizuki, N., Brusslan, J. A., Larkin, R., Nagatani, A. & Chory, J. Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc. Natl. Acad. Sci. United States Am.* **98**, 2053–8, DOI: 10.1073/pnas.98.4.2053 (2001).
11. Larkin, R. M., Alonso, J. M., Ecker, J. R. & Chory, J. GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Sci. (New York, N.Y.)* **299**, 902–6, DOI: 10.1126/science.1079978 (2003).
12. Jensen, P. E., Gibson, L. C., Henningsen, K. W. & Hunter, C. N. Expression of the chlI, chlD, and chlH genes from the Cyanobacterium synechocystis PCC6803 in Escherichia coli and demonstration that the three cognate proteins are required for magnesium-protoporphyrin chelatase activity. *The J. Biol. Chem.* **271**, 16662–16667, DOI: 10.1074/jbc.271.28.16662 (1996).
13. Reid, J. D., Siebert, C. A., Bullough, P. A. & Hunter, C. N. The ATPase activity of the ChlI subunit of magnesium chelatase and formation of a heptameric AAA+ ring. *Biochemistry* **42**, 6912–20, DOI: 10.1021/bi034082q (2003).
14. Reid, J. & Hunter, C. Magnesium-dependent ATPase activity and cooperativity of magnesium chelatase from Synechocystis sp. PCC6803. *J. Biol. Chem.* **279**, 26893, DOI: 10.1074/jbc.M400958200 (2004).
15. Viney, J., Davison, P. A., Hunter, C. N. & Reid, J. D. Direct Measurement of Metal-Ion Chelation in the Active Site of the AAA+ATPase Magnesium Chelatase. *Biochemistry* **46**, 12788–12794, DOI: 10.1021/bi701515y (2007).

16. Adams, N. B. P. & Reid, J. D. The Allosteric Role of the AAA+ Domain of ChlD Protein from the Magnesium Chelatase of Synechocystis Species PCC 6803. *J. Biol. Chem.* **288**, 28727–28732, DOI: 10.1074/jbc.M113.477943 (2013).

17. Adams, N. B. P. *et al.* Structural and functional consequences of removing the N-terminal domain from the magnesium chelatase ChlH subunit of Thermosynechococcus elongatus. *Biochem. J.* **464**, 315–322, DOI: 10.1042/BJ20140463 (2014).

18. Brindley, A. A., Adams, N. B. P., Hunter, C. N. & Reid, J. D. Five Glutamic Acid Residues in the C-Terminal Domain of the ChlD Subunit Play a Major Role in Conferring Mg 2+ Cooperativity upon Magnesium Chelatase. *Biochemistry* **54**, 6659–6662, DOI: 10.1021/acs.biochem.5b01080 (2015).

19. Adams, N. B. P., Brindley, A. A., Hunter, C. N. & Reid, J. D. The catalytic power of magnesium chelatase: a benchmark for the AAA+ATPases. *FEBS Lett.* **590**, 1687–1693, DOI: 10.1002/1873-3468.12214 (2016).

20. Farmer, D. A. *et al.* The ChlD subunit links the motor and porphyrin binding subunits of magnesium chelatase. *Biochem. J.* **476**, 1875–1887, DOI: 10.1042/BCJ20190095 (2019).

21. Reid, J. D. & Hunter, C. N. Current understanding of the function of magnesium chelatase. *Biochem. Soc. Transactions* **30**, 643–645, DOI: 10.1042/bst0300643 (2002).

22. Qian, P. *et al.* Structure of the Cyanobacterial Magnesium Chelatase H Subunit Determined by Single Particle Reconstruction and Small-angle X-ray Scattering. *J. Biol. Chem.* **287**, 4946–4956, DOI: 10.1074/jbc.M111.308239 (2012).

23. Chen, X. *et al.* Crystal structure of the catalytic subunit of magnesium chelatase. *Nat. Plants* **1**, 15125, DOI: 10.1038/nplants.2015.125 (2015).

24. Chen, Z., Zhang, X., Liu, Y. & Jiang, L. Crystal structure of the Arabidopsis thaliana C-terminal Chlh at 1.25A, DOI: 10.2210/pdb5ewu/pdb (2016).

25. Trott, O. & Olson, A. J. Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461, DOI: 10.1002/jcc.21334 (2010).

26. Karger, G. A., Reid, J. D. & Hunter, C. N. Characterization of the binding of deuteroporphyrin IX to the magnesium chelatase H subunit and spectroscopic properties of the complex. *Biochemistry* **40**, 9291–9, DOI: 10.1021/bi010562a (2001).

27. Porter, C. M. & Miller, B. G. Cooperativity in monomeric enzymes with single ligand-binding sites. *Bioorganic Chem.* **43**, 44 – 50, DOI: https://doi.org/10.1016/j.bioorg.2011.11.001 (2012).

28. Adams, N. B. P., Marklew, C. J., Brindley, A. A., Hunter, C. N. & Reid, J. D. Characterization of the magnesium chelatase from Thermosynechococcus elongatus. *Biochem. J.* **457**, 163–170, DOI: 10.1042/bj20130834 (2014).

29. Cornish-Bowden, A. *Fundamentals of Enzyme Kinetics* (Portland Press Ltd, 2004).

30. Schubert, H. L., Raux, E., Wilson, K. S. & Warren, M. J. Common Chelatase Design in the Branched Tetrapyrrole Pathways of Heme and Anaerobic Cobalamin Synthesis†. *Biochemistry* **38**, 10660–10669, DOI: 10.1021/bi9906773 (1999).

31. Al-Karadaghi, S., Hansson, M., Nikonov, S., Jönsson, B. & Hederstedt, L. Crystal structure of ferrochelatase: the terminal enzyme in heme biosynthesis. *Structure* **5**, 1501–1510, DOI: 10.1016/s0969-2126(97)00299-2 (1997).

32. Wu, C.-K. *et al.* The 2.0 A structure of human ferrochelatase, the terminal enzyme of heme biosynthesis. *Nat. Struct. Biol.* **8**, 156–160, DOI: 10.1038/84152 (2001).

33. Karlberg, T. *et al.* Metal Binding to *Saccharomyces cerevisiae* Ferrochelatase. *Biochemistry* **41**, 13499–13506, DOI: 10.1021/bi0260785 (2002).

34. Dailey, H. A. *et al.* Ferrochelatase at the millennium: Structures, mechanisms and [2Fe-2S] clusters. *Cell. Mol. Life Sci.* **57**, 1909–1926, DOI: 10.1007/PL00000672 (2000).

35. Dailey, H. A. *et al.* Prokaryotic heme biosynthesis: Multiple pathways to a common essential product. *Microbiol. Mol. Biol. Rev.* **81**, DOI: 10.1128/MMBR.00048-16 (2017).

36. Shen, Y. & Ryde, U. Reaction mechanism of porphyrin metallation studied by theoretical methods. *Chem. (Weinheim an der Bergstrasse, Ger.* **11**, 1549–64, DOI: 10.1002/chem.200400298 (2005).

37. Chen, G. E. *et al.* Complete enzyme set for chlorophyll biosynthesis in Escherichia coli. *Sci. Adv.* **4**, DOI: 10.1126/sciadv.aaq1407 (2018).

38. Smith, K. M. & Goff, D. A. Syntheses of Some Proposed Biosynthetic Precursors to the Isocyclic Ring in Chlorophyll a. *J. Org. Chem.* **51**, 657–666, DOI: 10.1021/jo00355a015 (1986).

39. Jensen, P. E., Gibson, L. C. D. & Hunter, C. N. Determinants of catalytic activity with the use of purified I, D and H subunits of the magnesium protoporphyrin IX chelatase from Synechocystis PCC6803. *Biochem. J.* **334**, 335–344, DOI: 10.1042/bj3340335 (1998).

40. Sawicki, A. & Willows, R. D. Kinetic analyses of the magnesium chelatase provide insights into the mechanism, structure, and formation of the complex. *The J. biological chemistry* **283**, 31294–302, DOI: 10.1074/jbc.M805792200 (2008).

41. Celis, A. I. & DuBois, J. L. Making and breaking heme. *Curr. Opin. Struct. Biol.* **59**, 19–28, DOI: 10.1016/j.sbi.2019.01.006 (2019).

42. Winter, G. *xia2*: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190, DOI: 10.1107/S0021889809045701 (2010).

43. Sheldrick, G. M. & Schneider, T. R. Shelxl: High-resolution refinement. In *Macromolecular Crystallography Part B*, vol. 277 of *Methods in Enzymology*, 319–343, DOI: https://doi.org/10.1016/S0076-6879(97)77018-6 (Academic Press, 1997).

44. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. Sect. D* **66**, 486–501, DOI: 10.1107/S0907444910007493 (2010).

45. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta crystallographica. Sect. D, Biol. crystallography* **62**, 1002–1011, DOI: 10.1107/S0907444906022116 (2006).

46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D* **53**, 240–255, DOI: 10.1107/S0907444996012255 (1997).

47. Chen, V. B. *et al. MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D* **66**, 12–21, DOI: 10.1107/S0907444909042073 (2010).

48. Winn, M. D. *et al.* Overview of the *CCP*4 suite and current developments. *Acta Crystallogr. Sect. D* **67**, 235–242, DOI: 10.1107/S0907444910045749 (2011).

49. Nicholls, R. A., Long, F. & Murshudov, G. N. Low-resolution refinement tools in *REFMAC*5. *Acta Crystallogr. Sect. D* **68**, 404–417, DOI: 10.1107/S090744491105606X (2012).

50. Nicholls, R. A., Kovalevskiy, O. & Murshudov, G. N. *Low Resolution Refinement of Atomic Models Against Crystallographic Data*, 565–593 (Springer New York, New York, NY, 2017).

51. O'Boyle, N. M. *et al.* Open Babel: An Open chemical toolbox. *J. Cheminformatics* **3**, 1–14, DOI: 10.1186/1758-2946-3-33 (2011).

52. Forli, S. *et al.* Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **11**, 905–919, DOI: 10.1038/nprot.2016.051 (2016).

53. Walser, M. Ion association v. dissociation constants for complexes of citrate with sodium, potassium, calcium, and magnesium ions. *The J. Phys. Chem.* **65**, 159–161, DOI: 10.1021/j100819a045 (1961).

54. Meneely, K. M., Sundlov, J. A., Gulick, A. M., Moran, G. R. & Lamb, A. L. An Open and Shut Case: The Interaction of Magnesium with MST Enzymes. *J. Am. Chem. Soc.* **138**, 9277–9293, DOI: 10.1021/jacs.6b05134 (2016).

55. Stein, R. L. *Kinetics of Enzyme Action* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011).

Extended Data

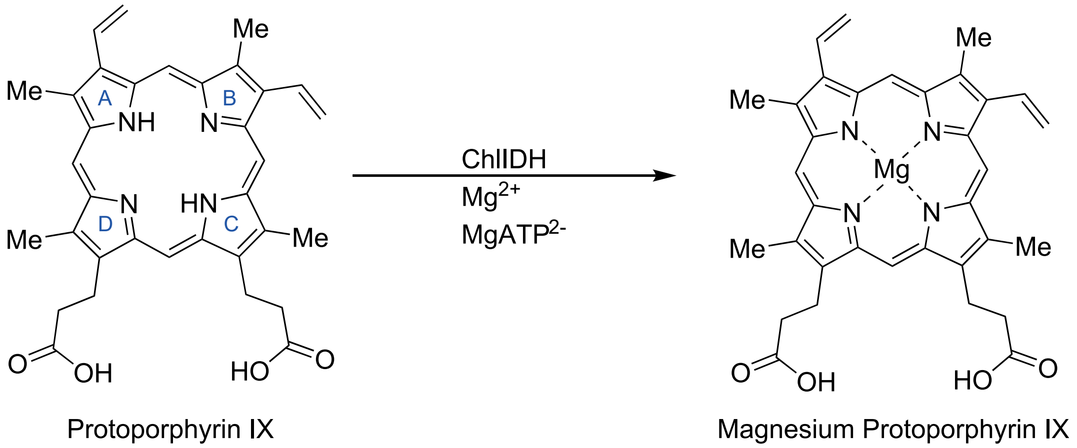

Figure E1. Magnesium chelatase (ChlIDH) catalyses the ATP dependent conversion of protoporphyrin IX (P_{IX}) into magnesium protoporphyrin IX (MgP_{IX}). The pyrrole rings are identified by the letter in blue.

Data Collection	SeMet SAD T_ChlH_926	Native T_ChlH_926 (PDB:6YS9)	WT Syn_ChlH (PDB:6YSG)	E660D Syn_ChlH (PDB:6YT0)	E660W Syn_ChlH (PDB:6YTN)	E625K Syn_ChlH (PDB:6YTJ)
Beamline	DLS, i24 (mx8987-28)	DLS, i03 (mx8987-42)	DLS, i02 (mx8987-45)	DLS, i04 (mx12788-25)	DLS, i03 (mx12788-5)	DLS, i03 (mx12788-5)
Wavelength (Å)	0.97951	0.97633	0.98674	0.97626	0.97630	0.97630
Resolution (Å)	65.75-1.84 (1.89-1.84)	64.38-1.64 (1.68-1.64)	83.78-2.54 (2.74-2.54)	54.6-2.85 (2.90-2.85)	79.39-2.7 (2.77-2.7)	80.11-2.79 (2.86-2.79)
Space Group	$P2_1$	$P2_1$	H3	H3	H3	H3
Unit Cell Dimensions (a, b, c, α, β, γ)	76.48, 140.6, 78.59, 90, 108.82, 90	76.48, 140.33, 78.01, 90, 108.66, 90	321.12, 321.12, 104.97, 90, 90, 120	321.79, 321.97, 104.92, 90, 90, 120	317.55, 317.55, 104.43, 90, 90, 120	320.42, 320.42, 104.88, 90, 90, 120
Total reflections[c]	850868 (38925)	395673 (23118)	1143007 (227759)	547095 (26068)	627443 (47821)	570782 (42763)
Unique reflections[c]	134717 (9251)	172739 (10643)	133090 (27461)	94580(4748)	107741 (7995)	99929 (7451)
Multiplicity[c]	6.3 (4.2)	2.3 (2.2)	8.6 (8.3)	5.8 (5.5)	5.8 (6.0)	5.7 (5.7)
Completeness[c] (%)	99.2 (92.5)	91.2 (76.3)	100 (100)	100 (100)	99.9 (100)	100 (100)
Mean I/σ [c] (I)	11.9 (1.6)	7.4 (1.6)	12.1 (1.7)	9.4 (1.5)	10.2 (1.3)	6.4 (1.2)
R_{pim}[b,c]	0.040 (0.498)	0.057 (0.442)	0.041 (0.495)	0.061 (0.544)	0.059 (0.676)	0.074 (0.659)
CC-half	0.998 (0.482)	0.994 (0.487)	0.999 (0.638)	0.995 (0.560)	0.997 (0.490)	0.993 (0.583)
Anomalous slope	0.923					
Anomalous multiplicity	3.1 (2.2)					
Anomalous completeness	95.2 (80)					
Refinement						
R_{factor}		0.20	0.20	0.21	0.19	0.21
R_{free}		0.25	0.25	0.26	0.23	0.27
RMSD bonds (Å)		0.0078	0.0072	0.0075	0.0041	0.0051
RMSD angles (°)		1.55	1.55	1.60	1.29	1.38
No. of non-H atoms						
Protein		12024	19588	19321	19763	19393
Ligands		4	13	0	0	0
Water		553	114	32	8	19
Average B factors						
Main chain		22.3	61.4	63.3	74.0	59.3
Side chains		31.6	68.1	66.2	76.9	74.2
Ligands		28.8	93.4	-	-	-
Water		23.7	38.9	30.9	47.2	43.2
Ramachandran favored/allowed (%)		96.7/3.2	93.9/5.6	93.4/6.0	94.7/4.8	92.7/6.7
Molprobity Score		1.29 (97th percentile; N=7200, 1.61Å± 0.25Å)	2.02 (95th percentile; N=6647, 2.54Å± 0.25Å)	2.31 (96th percentile; N=4562, 2.85Å± 0.25Å)	1.85 (99th percentile; N=5412, 2.7Å± 0.25Å)	2.12 (98th percentile; N=4482, 2.79Å± 0.25Å)

Table E1. [a] $R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$. [b] $R_{pim} = \Sigma_{hkl} \sqrt{1/n\text{-}1}\Sigma_{i=1} \mid I_i - I_m \mid / \Sigma_{hkl}\Sigma_i I_i$, where I_i and I_m are the observed intensity and mean intensity of related reflections, respectively. [c] Values in parenthesis are for data in the high-resolution shell.

Protein	Binding site	$k_{\mathrm{cat}}{}^{\mathrm{D_{IX}}}$ / min^{-1}	$K_{\mathrm{M}}{}^{\mathrm{D_{IX}}}$ / μM	$n^{\mathrm{D_{IX}}}$	$k_{\mathrm{cat}}/K_{\mathrm{M}}{}^{\mathrm{D_{IX}}}$ / min^{-1} μM^{-1}	$K_{\mathrm{d}}{}^{\mathrm{D_{IX}}}$ / μM
WT	N/A	0.56 ± 0.01	1.80 ± 0.13		0.31 ± 0.02	0.65 ± 0.38
T987A†	1	0.39 ± 0.01	1.63 ± 0.12	1.9 ± 0.2	0.24 ± 0.02	2.21 ± 0.46
D988A	1	0.27 ± 0.01	3.49 ± 0.43		0.08 ± 0.01	6.32 ± 1.76
S1039A†	1	0.44 ± 0.05	6.80 ± 1.19	1.7 ± 0.2	0.06 ± 0.01	5.91 ± 4.21
V1041A†	1	0.41 ± 0.02	2.63 ± 0.24	1.7 ± 0.2	0.16 ± 0.02	2.32 ± 1.29
R667A/K/E	CobN	n.d.	n.d.		n.d.	9.26 ± 5.18
K991A *	CobN	0.23 ± 0.03	11.92 ± 2.91		0.02 ± 0.01	4.53 ± 2.15
S1103A	2	0.24 ± 0.01	1.18 ± 0.15		0.20 ± 0.03	1.45 ± 0.49
K1129A	2	0.18 ± 0.02	4.00 ± 0.96		0.04 ± 0.01	10.5 ± 5.4
D1177A$^{\triangledown}$	2	0.37 ± 0.02	0.48 ± 0.10		0.77 ± 0.17	0.91 ± 0.23
S1178A	2	0.17 ± 0.01	0.8 ± 0.10		0.21 ± 0.02	0.23 ± 0.20

Table E2. Fit coefficients for mutations in the active site of the ChlH protein with respect to [$\mathrm{D_{IX}}$]. Data from Fig. 3 panels **d** and **e** is described by equation 5 unless stated otherwise, errors reported as ± one standard deviation of the fit coefficient. n.d. no activity detected. * these residues are adjacent to site 1. † Fitted to the Hill equation (6), where $s_{0.5}$ is reported instead of K_{M}, and $k_{\mathrm{cat}}/s_{0.5}$ instead of $k_{\mathrm{cat}}/K_{\mathrm{M}}$. $^{\triangledown}$ Fitted to the substrate inhibition equation (7), where $K_{\mathrm{i}} = 37.1 \pm 7.7$ μM.

Protein	k_{cat} / min^{-1}		K_{M}*		$k_{\mathrm{cat}}/K_{\mathrm{M}}{}^{\dagger}$		n
	$\mathrm{D_{IX}}$	$\mathrm{Mg^{2+}}$	$\mathrm{D_{IX}}$ / μM	$\mathrm{Mg^{2+}}$ / mM	$\mathrm{D_{IX}}$ / $\mu\mathrm{M}^{-1}\mathrm{min}^{-1}$	$\mathrm{Mg^{2+}}$ / $\mathrm{mM}^{-1}\mathrm{min}^{-1}$	$\mathrm{Mg^{2+}}$
WT	0.45±0.01	0.45± 0.03	1.42±0.12	2.32±0.21	0.32±0.03	0.19±0.02	1.7±0.3
H1174V	0.42±0.01	0.42±0.06	1.52±0.17	2.76±0.68	0.28±0.03	0.15±0.04	1.4± 0.3
E660D	0.11±0.007	0.065±0.005	14.49±1.97	1.47±0.18	0.008±0.040	0.044±0.006	1.6±0.5

Table E3. Characteristic steady state parameters of WT ChlH, H1174V and E660D proteins. Data from Fig. E2. When magnesium is the varied substrate, data is fitted to Equation 6: * $S_{0.5}$, † k_{cat} / $S_{0.5}$. Values reported are the errors reported as ± one standard deviation of the fit coefficient.

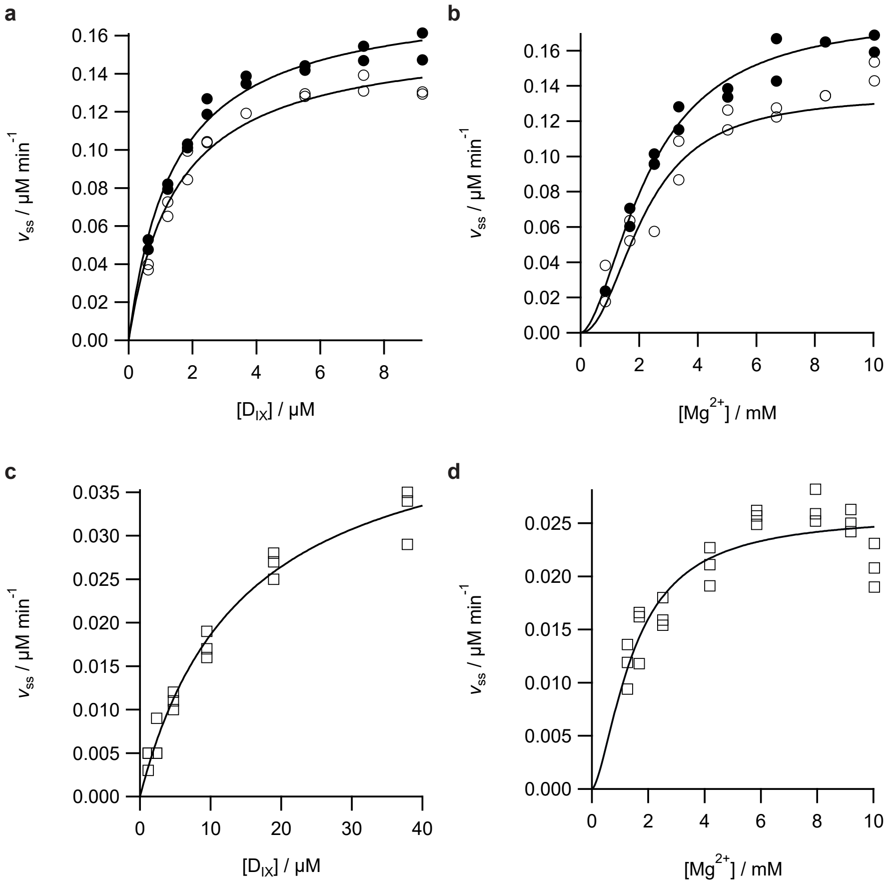

Figure E2. **a** and **c**, D_{IX} and **b** and **d**, Mg^{2+} dependence of the steady state rate of Mg^{2+} chelation for WT ChlH (closed circles), H1174V (open circles), E660D (open squares). Assays contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT ChlH or H1174V in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 5 mm $MgATP^{2-}$ unless stated otherwise. Lines are theoretical, where steady state rates (v_{ss}) were fitted to the Michaelis-Menten equation (Equation 5) in panel **a** and **c** or the Hill equation (Equation 6) in panel **b** and **d**, and kinetic coefficients reported in Table E3. Each data point is an individual experiment

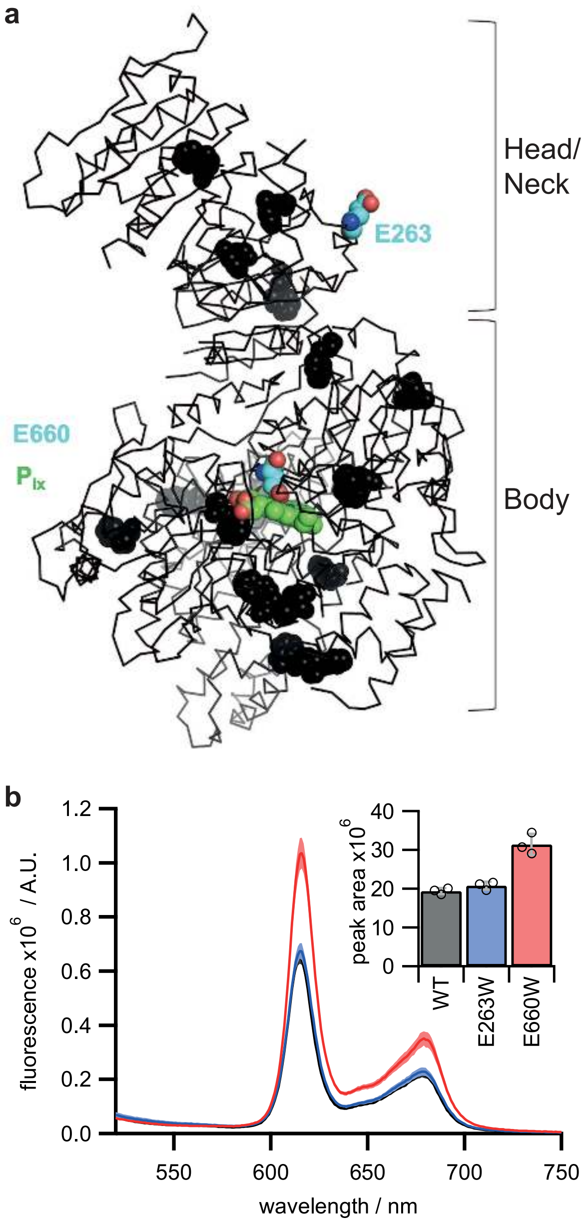

Figure E3. Energy transfer from tryptophans to porphyrin in the active site. **a**, The structure of ChlH represented as black ribbon with all native tryptophan residues represented as black spheres. The control mutation, E263, is shown as cyan spheres, as is the active site E660, with P_{IX} docked in the active site represented as green spheres. **b**, Fluorescence emission of D_{IX} after excitation or tryptophans at 280 nM. Black, WT ChlH; blue, E263W, Red, E660W. Averages of three independent biological repeats are represented with a central solid line and shading representing standard deviation. The peak areas in the insert are the total integrated number of counts for the full wavelength range shown in the graph. The bars represent the mean value from three independent experiments, with error bars ± the standard deviation. In addition each area value is plotted for as an open circle.

Supplementary Data

Table S1. Summary of all mutants produced in this work for kinetic analysis. When magnesium is the varied substrate, data is fitted to Equation 6: * $S_{0.5}$, † k_{cat} / $S_{0.5}$. n.d. not determined, – no chelatase activity observed. ⋆ these residues are adjacent to site 1. ‡ Fitted to the Hill equation (6), where $s_{0.5}$ is reported instead of K_M, and $k_{cat}/s_{0.5}$ instead of k_{cat}/K_M. ▽ Fitted to the substrate inhibition equation (7), where K_i = 37.1 ± 7.7 μM. All numerical values are the reported fit coefficients from non-linear regression analysis, with errors reported as ± one standard deviation.

Protein	**Area**	**Binding Site**	**Active**	**k_{cat} / min^{-1}**		**K_M***		**k_{cat}/K_M†**		***n***		**K_d^{DIX}**	**Figure Ref**
				D_{IX}	Mg^{2+}	D_{IX} / μM	Mg^{2+} / mM	D_{IX} / $\mu M^{-1}min^{-1}$	Mg^{2+} / $mM^{-1}min^{-1}$	D_{IX}	Mg^{2+}		
WT				0.45 ± 0.01	0.45 ± 0.03	1.42 ± 0.12	2.23 ± 0.21	0.32 ± 0.03	0.19 ± 0.02		1.7 ± 0.3	0.65 ± 0.38	3,E2,S4
T987A‡	Central Cavity	1	Y	0.39 ± 0.01	n.d.	1.63 ± 0.12	n.d.	0.24 ± 0.02	n.d.	1.9 ± 0.2	n.d.	2.21 ± 0.46	3,S4
D988A	Central Cavity	1	Y	0.27 ± 0.01	n.d.	3.49 ± 0.43	n.d.	0.08 ± 0.01	n.d.		n.d.	6.32 ± 1.76	3,S4
S1039A‡	Central Cavity	1	Y	0.44 ± 0.05	n.d.	6.80 ± 1.19	n.d.	0.06 ± 0.01	n.d.	1.7 ± 0.2	n.d.	5.91 ± 4.21	3,S4
V1041A‡	Central Cavity	1	Y	0.41 ± 0.02	n.d.	2.63 ± 0.24	n.d.	0.16 ± 0.02	n.d.	1.7 ± 0.2	n.d.	2.32 ± 1.29	3,S4
R667A/K/E	Surface Exposed	CobN	N	–	–	–	–	–	–	–	–	9.26 ± 5.18	3,S4
K991A⋆	Central Cavity	CobN	Y	0.23 ± 0.03	n.d.	11.92 ± 2.91	n.d.	0.02 ± 0.01	n.d.		n.d.	4.53 ± 2.15	3,S4
S1103A	Central Cavity	2	Y	0.24 ± 0.01	n.d.	1.18 ± 0.15	n.d.	0.20 ± 0.03	n.d.		n.d.	1.45 ± 0.49	3,S4
K1129A	Central Cavity	2	Y	0.18 ± 0.02	n.d.	4.00 ± 0.96	n.d.	0.04 ± 0.01	n.d.		n.d.	10.5 + 5.4	3,S4
D1177A▽	Central Cavity	2	Y	0.37 ± 0.02	n.d.	0.48 ± 0.10	n.d.	0.77 ± 0.17	n.d.		n.d.	0.91 ± 0.23	3,S4
S1178A	Central Cavity	2	Y	0.17 ± 0.01	n.d.	0.80 ± 0.10	n.d.	0.21 ± 0.02	n.d.		n.d.	0.23 ± 0.20	3,S4
H1174V	Catalytic site		Y	0.42 ± 0.01	0.42 ± 0.06	1.52 ± 0.17	2.76 ± 0.68	0.28 ± 0.03	0.15 ± 0.04		1.4 ± 0.3	0.77 ± 0.49	E2,S6
E660D	Catalytic site		Y	0.11 ± 0.007	0.065 ± 0.005	14.49 ± 1.97	1.47 ± 0.18	0.008 ± 0.040	0.044 ± 0.006		1.6 ± 0.5	3.05 ± 0.53	E2,S6
E660W	Catalytic site		N	–	–	–	–	–	–	–	–	0.30 ± 0.01	E3,S7
E660R/H/Q	Catalytic site		N	–	–	–	–	–	–	–	–	n.d.	S8
E625D	Buried Channel		N	–	–	–	–	–	–	–	–	0.91 ± 0.40	4,S10
E625K/H/Q	Buried Channel		N	–	–	–	–	–	–	–	–	n.d.	4
Y653T	Buried Channel		Y	n.d.	0.16 ± 0.01	n.d.	4.41 ± 0.38	n.d.	0.036 ± 0.004		2.4 ± 0.3	2.15 ± 1.40	4,S10

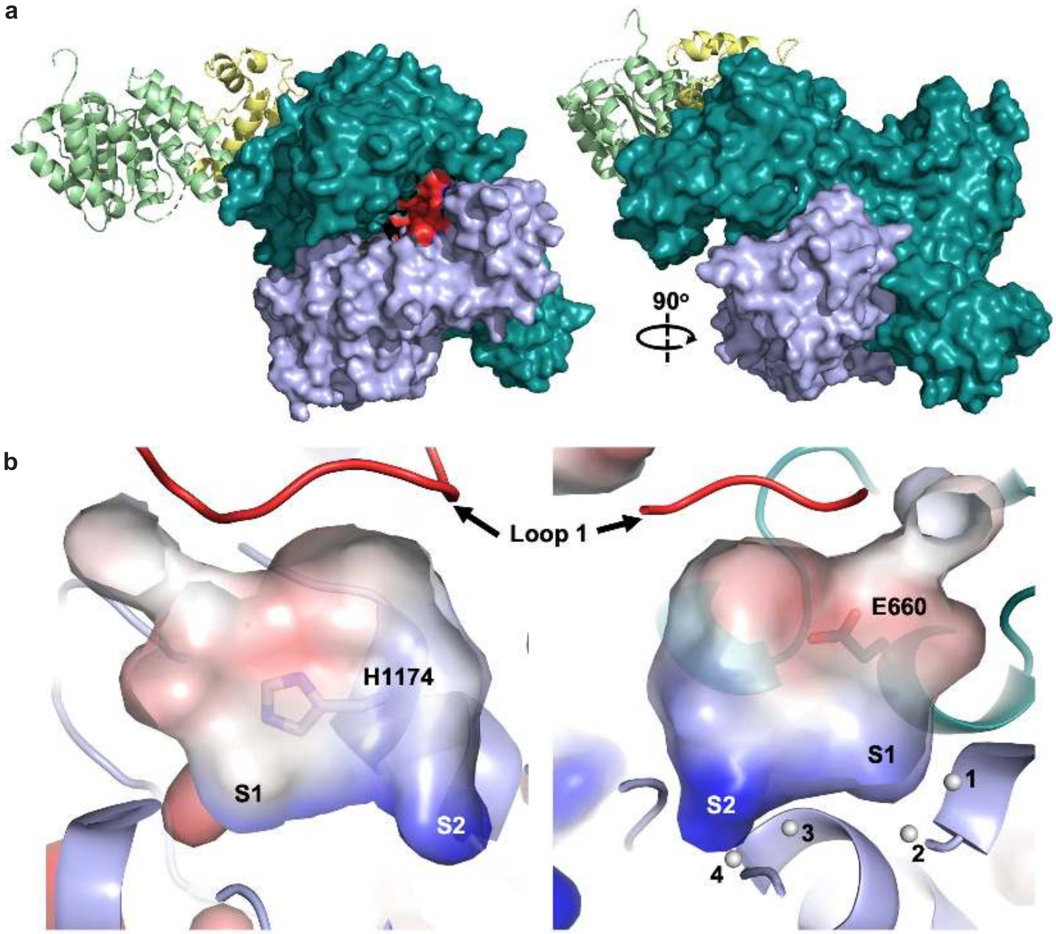

Figure S1. Relationship between the two sub-domains within the body of Syn_ChlH and variation in surface charge inside the active site cavity. **a**, Perpendicular views of the two sub-domains of the body of Syn_ChlH are shown with a solid surface and coloured according to the C-terminal portion (Body II; pale blue), which is equivalent to T_ChlH_926, and the N-terminal portion (Body I; teal), which forms a C-shaped domain. The head and neck domains are coloured green and yellow, respectively and are shown as cartoons. Loop 1 is coloured red, indicating the position of the active site cavity, buried between the two domains. **b**, An open book view of the electrostatics of the inside surface of the active site cavity (red=negative, white=neutral, blue=positive). The four mainchain amides sat at the base of the pocket are labelled 1-4 ((1=T987, 2=G986, 3=V1041, 4=G1040) and the relative positions of site 1 and 2 and indicated. Loop 1 is shown in red.

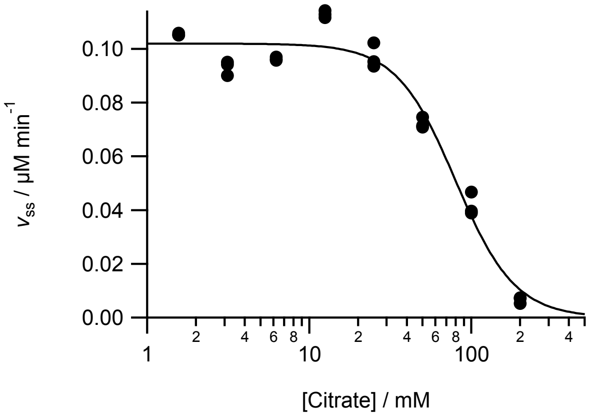

Figure S2. Magnesium chelatase is inhibited by citrate. A standard magnesium chelatase assay was spiked with citrate (from 200 - 1.6 mM) and the steady state rate plotted against citrate concentration, displaying inhibition behaviour, with a calculated $K_{\mathrm{i,app}}$ of 80 mM. Assays contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT or mutant ChlH in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 8 μM D_{IX} 5 mm $MgATP^{2-}$, pH 7.7. As citrate binds to magnesium ions in solution[53], the assays contained an excess of 10 mM $MgCl_2$ over citrate to ensure there were free magnesium ions present in solution for the reaction to proceed.

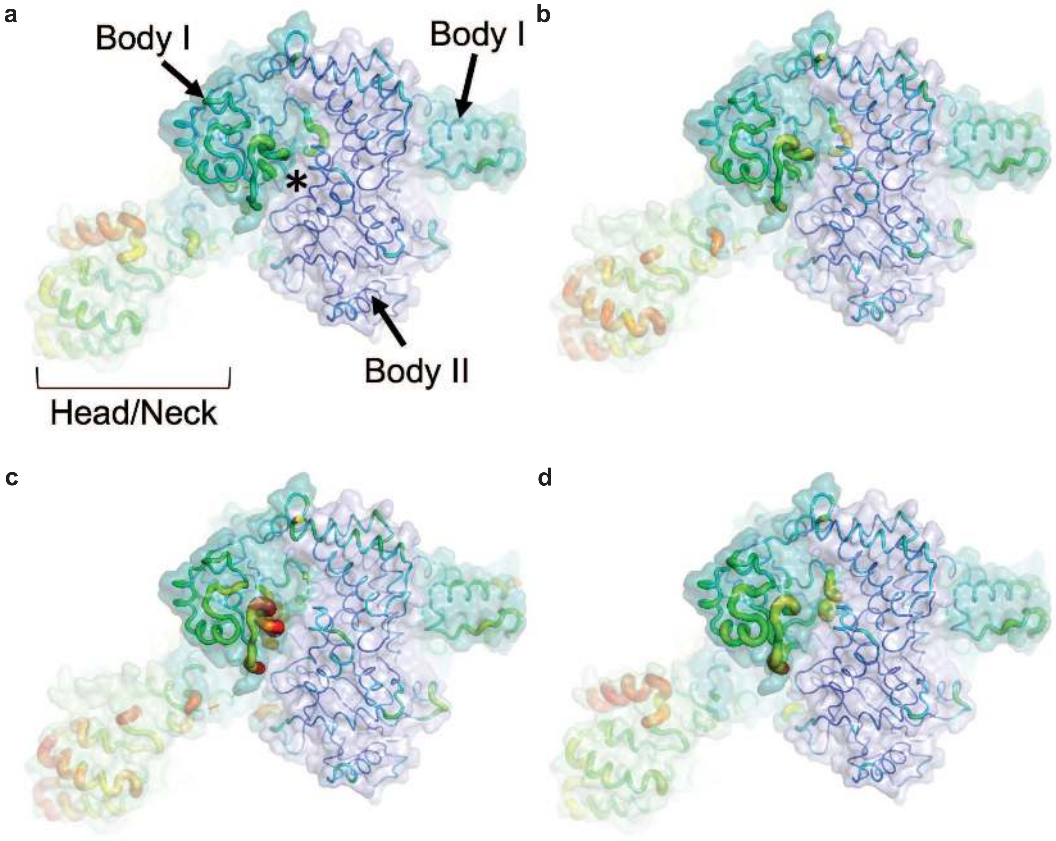

Figure S3. B-factor putty representations of WT and mutant Syn_ChlH subunits showing flexibility in the region of the Body I domain adjacent to the active site. **a**, The WT structure of Syn_ChlH is drawn as B-factor putty from low B-factors (blue/thin tubes) to higher B-factors (red/fat tubes). A partially transparent surface showing each of the domains of the protein is overlaid, focusing on the two body domains (Body I in teal and Body II in pale blue). B-factors are generally lower in the Body II domain, showing the protein backbone is more rigid. In the region of Body I adjacent to the active site (*), where the Loops 1-3 are, the B-factors are on average higher, showing this region is more flexible. Parts **b-d** show the E660D, E625K and E660W mutants, respectively. In all structures, residues in the head and neck (H/N) domains have the highest overall B-factors.

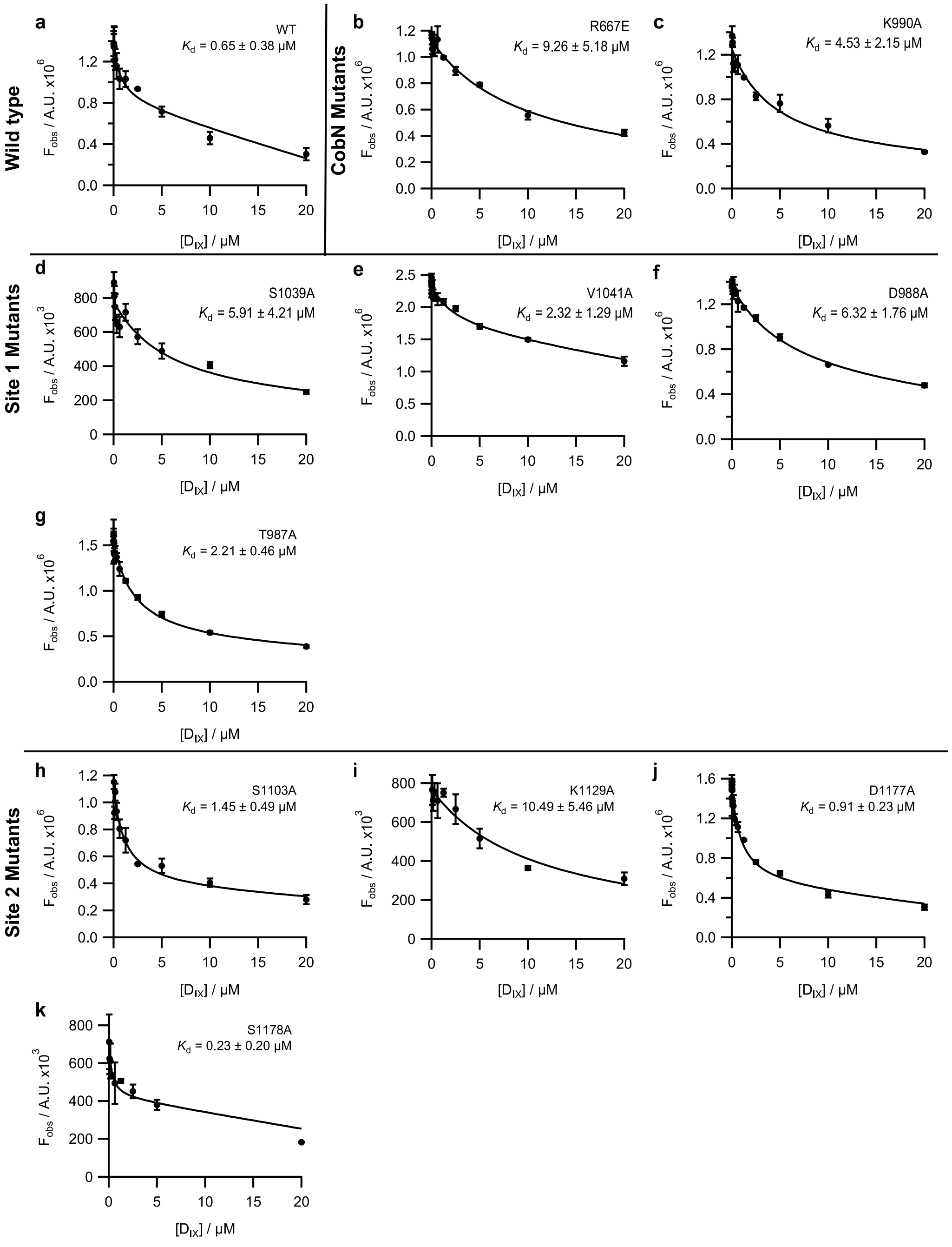

Figure S4. Binding studies of D_{IX} to ChlH mutants. Dissociation constants (K_d) of D_{IX} binding to ChlH WT and mutants were determined by monitoring tryptophan fluorescence quenching. Experiments were performed in assay buffer 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, pH 7.7 at 34 °C. **a - k**, Binding isotherms for WT, CobN mutants, site 1 and site 2 mutants. K_ds are reported next to each protein, and in table Table E2. Error bars indicate the standard deviation from the mean of three independent titrations

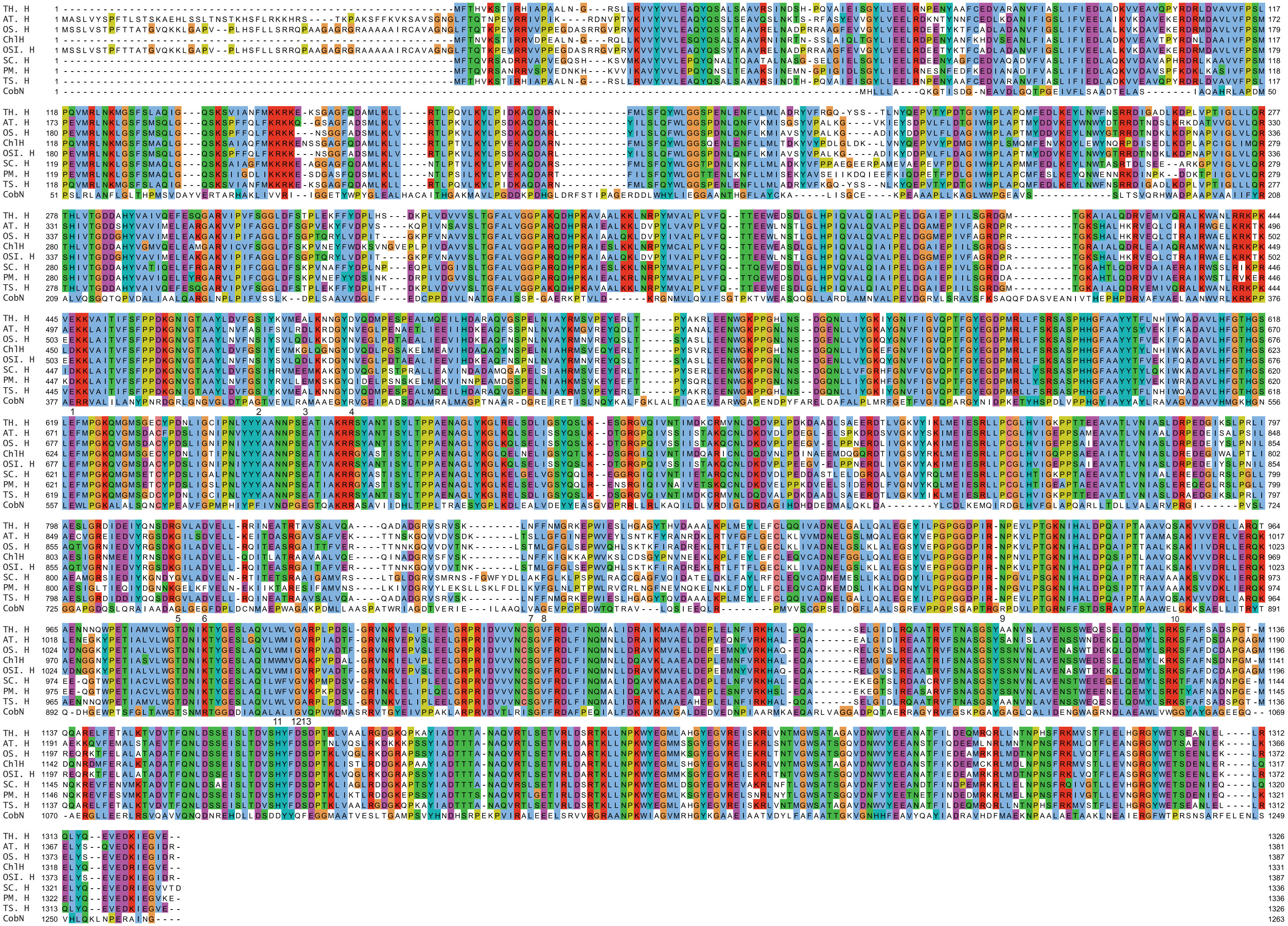

Figure S5. Sequence alignment of ChlH and CobN proteins. TH. H, *Thermosynecococcus elongatus*; AT. H, *Arabidopsis thaliana*; OS. H, *Oryza sativa subsp. japonica*; ChlH, *Synechocystisis sp. 6803*; SC. H, *Synechococcus sp. (strain WH8102)*; PM. H, *Prochlorococcus marinus (strain MIT 9301)*; TS. H, *Thermosynechococcus sp. NK55a*; CobN, *Brucella melitensis*. Residues investigated in this study: 1, E625; 2, Y653; 3, E660; 4, R667; 5, T987; 6, K991; 7, S1039; 8, V1041; 9, S1103; 10, K1129; 11, H1174; 12, D1177; 13, S1178.

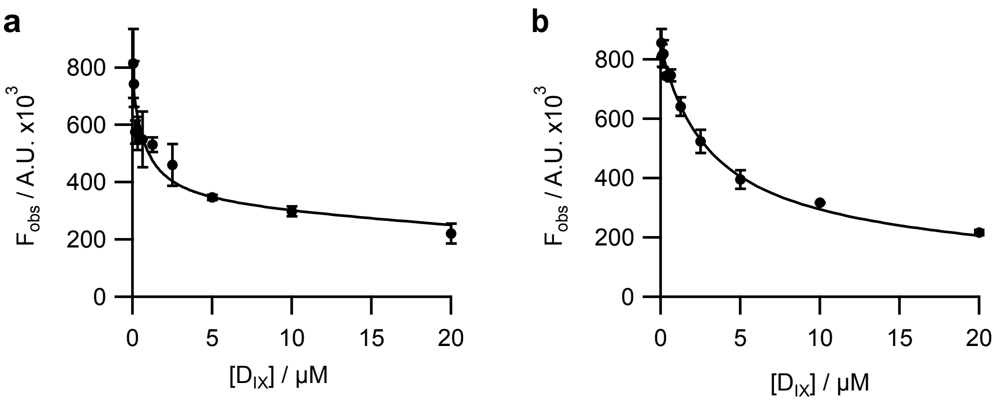

Figure S6. Binding of D_{IX} to ChlH H1174V and E660D. **a**, ChlH H1174V has a similar K_d for D_{IX} as WT protein, K_d= 0.77 ± 0.49 μM, where as **b**, E660D has a K_d = 3.05 ± 0.53 μM. Error bars indicate the standard deviation from the mean of three independent titrations.

Table S2. Primers used for subcloning *Thermosynechoccus elongatus* T_926 truncation

Protein	Mutation/Cut site	Primer forward	Primer reverse
T_ChlH	T926 truncation	ATCCAATTCATATGCCTGAGGTTTTGC	AAAACCTCAGGCATATGAATTGGATCG

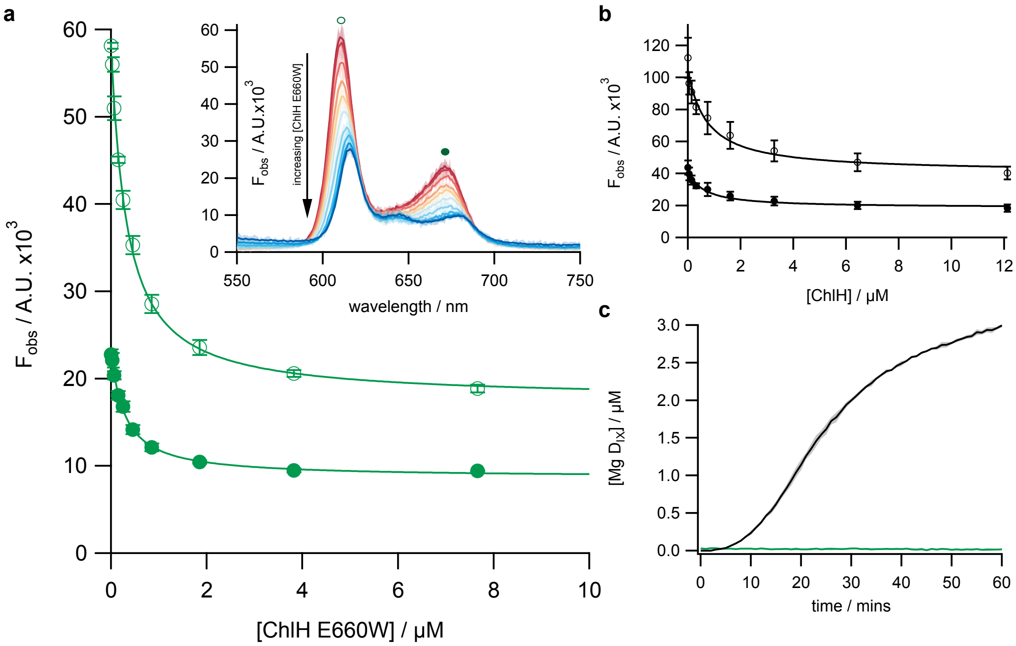

Figure S7. Binding and activity data of the ChlH E660W mutant. **a** Dissociation constant (K_d) of D_{IX} to E660W was determined by monitoring D_{IX} fluorescence quenching with increasing concentration of protein (see insert). Fluorescent intensity at 611 nm (open markers) and 672 nm (closed markers) are plotted, K_d = 0.30 ± 0.01 μM. This 'enzyme in excess' method of determining K_d for D_{IX} was internally consistent with a ChlH WT titration, panel **b**, K_d = 0.66 ± 0.12 μM (compare with K_d of 0.66 μM see Fig. S4 panel A). In panels a and b, data are presented as the mean of 3 repeats ± 1 standard deviation. **c** Magnesium chelatase activity of WT (black line) protein compared with E660W (green line) where no activity for the E600W protein is observed. Assays contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT or mutant ChlH in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 8 μM D_{IX}, 5 mm $MgATP^{2-}$, pH 7.7. For line graphs, the solid lines indicate the average of 3 independent experiments with the standard deviation shown in shading.

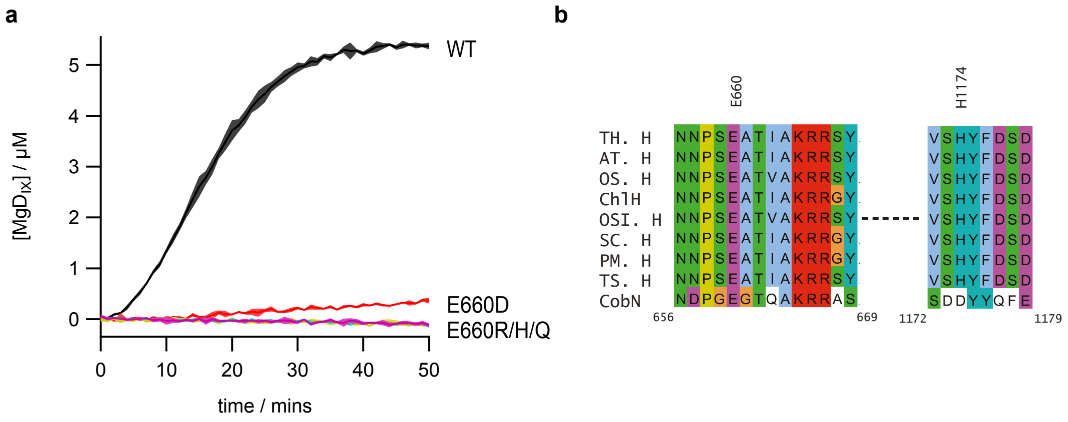

Figure S8. Activity of ChlH E660 mutants and sequence alignments highlighting conserved E660 and H1174 residues. **a**, Time courses of MgCH assays with WT, E625 and E660 mutants of ChlH; assays were performed in triplicate and are displayed as the mean with standard deviations shown in shading. Assays contained 0.1 μM ChlD, 0.2 μM ChlI, 0.4 μM WT or mutant ChlH in 50 mM MOPS/KOH, 0.3 M glycerol, 1 mM DTT, 10 mM free Mg^{2+}, 8 μM D_{IX}, 5 mm $MgATP^{2-}$, pH 7.7. **b**, Sequence alignment of ChlH and CobN highlighting conserved Glutamic Acid 660 and Histidine 1174 in *Synechocystistis*. TH. H, *Thermosynecoccus elongatus*; AT. H, *Arabidopsis thaliana*; OS. H, *Oryza sativa subsp. japonica*; ChlH, *Synechocystistis sp. 6803*; SC. H, *Synechococcus sp. (strain WH8102)*; PM. H, *Prochlorococcus marinus (strain MIT 9301)*; TS. H, *Thermosynechococcus sp. NK55a*; CobN, *Brucella melitensis*

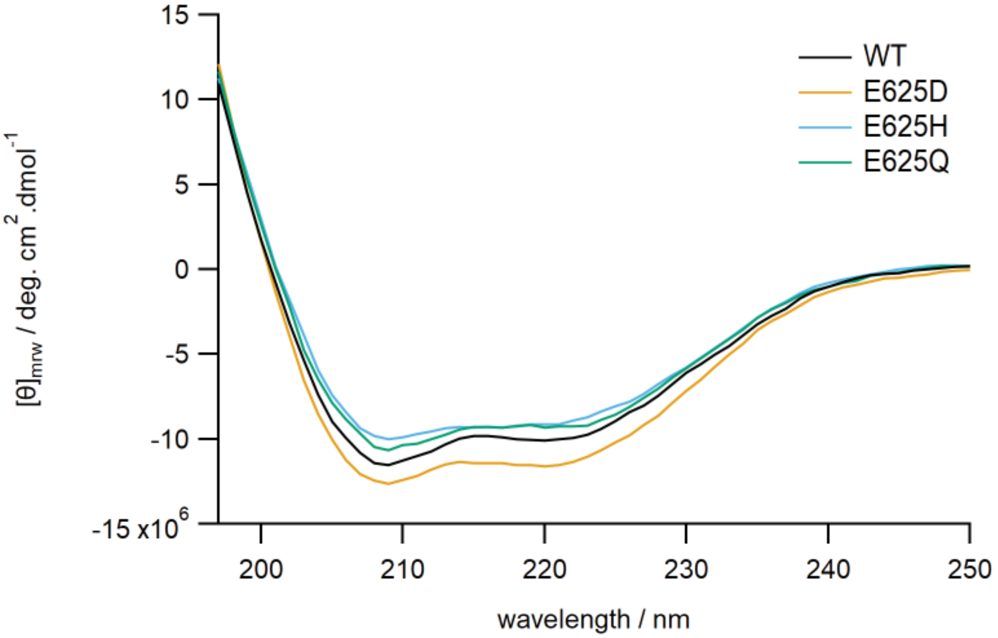

Figure S9. CD spectra of E625 mutants. E625 mutations do not have have gross changes in secondary structure. CD spectra (mean residue elipticity) measured at 25 ° in 5 mM sodium phosphate, pH 7.4, protein identity as indicated by legend in figure.

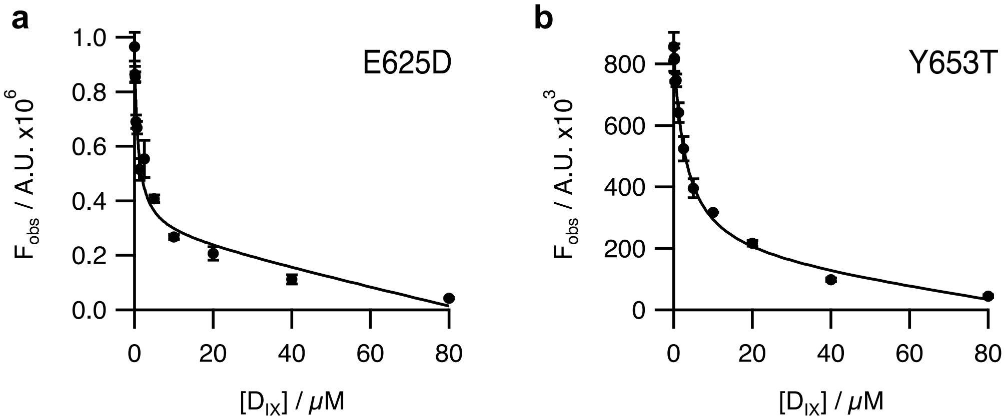

Figure S10. D_{IX} binding studies of ChlH E625D and Y653T mutants. Both E625D and Y653T mutations are able to bind D_{IX} to a similar degree as WT. **a**, E625D K_d=0.91±0.40 μM; **b**, Y653T, K_d=2.15±1.40 μM. Error bars indicate the standard deviation from the mean of three independent titrations

Table S3. Primers used for site directed mutagensis

Protein	Mutation	Primer forward	Primer reverse
ChlH	E625Q	CTGTTTACCGGGCATAAACTGCAAAGAACCGTGGGTGCC	GGCACCCACGGTTCTTTGCAGTTTATGCCCGGTAAACAG
ChlH	E625D	TGTTTACCGGGCATAAAATCCAAAGAACCGTGGGT	ACCCACGGTTCTTTGGATTTTATGCCCGGTAAACA
ChlH	E625H	CTGTTTACCGGGCATAAAATGCAAAGAACCGTGGGTGCC	GGCACCCACGGTTCTTTGCATTTTATGCCCGGTAAACAG
ChlH	E625K	GCACCCACGGTTCTTTGAAATTTATGCCCGGTAAA	TTTACCGGGCATAAATTTCAAAGAACCGTGGGTGC
ChlH	Y653F	CCATTCCCAACTTGTACTACTTCGCCGCCAACA	TGTTGGCGGCGAAGTAGTACAAGTTGGGAATGG
ChlH	S659A	ATGGTAGCTTCAGCAGGATTGTTGGCGGC	CCGCCAACAATCCTGCTGAAGCTACCATCG
ChlH	E660A	AACAATCCTTCTGCAGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTGCAGAAGGATTGTTGG
ChlH	E660W	AACAATCCTTCTTGGGCTACCATCGCCAAACG	TTTGGCGATGGTAGCCCAAGAAGGATTGTTGG
ChlH	E660D	AACAATCCTTCTGATGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTCGAGAAGGATTGTTGG
ChlH	E660H	AACAATCCTTCTCATGCTACCATCGCCAAACG	TTTGGCGATGGTAGCATGAGAAGGATTGTTGG
ChlH	E660Q	AACAATCCTTCTCAAGCTACCATCGCCAAACG	TTTGGCGATGGTAGCTTGAGAAGGATTGTTGG
ChlH	E660R	AACAATCCTTCTCGAGCTACCATCGCCAAACG	TTTGGCGATGGTAGCATCAGAAGGATTGTTGG
ChlH	R667A	TGGACGCATAACCCGCCCGTTTGGCGATGG	CCATCGCCAAACGGGCGGGTTATGCGTCCA
ChlH	R667K	GGTGGACGCATAACCCTTCCGTTTGGCGATGGTA	TACCATCGCCAAACGGAAGGGTTATGCGTCCACC
ChlH	R667E	GTGGACGCATAACCCTCCCGTTTGGCGATGGT	ACCATCGCCAAACGGGAGGGTTATGCGTCCAC
ChlH	D988K	TCTGTGCTCTGGGGTACGAAGAACATCAAAACCTACGGC	GCCGTAGGTTTTGATGTTCTTCGTACCCCAGAGCACAGA
ChlH	T987A	TTCTGTGCTCTGGGGTGCGGATAACATCAAAACC	TAGGTTTTGATGTTATCCGCACCCCAGAGCACAGAAGC
ChlH	D988A	TTCTGTGCTCTGGGGTACGGCTAACATCAAAACC	TTTTGATGTTAGCCGTACCCCAGAGCACAGAAGC
ChlH	K991A	TACGGATAACATCGCAACCTACGGCGAATCC	ATTCGCCGTAGGTTGCGATGTTATCCGTACC
ChlH	S1039A	TGGTGGTGAATTGTGCCGGGGTATTCC	ATCCCGGAATACCCCGGCACAATTCACCACC
ChlH	S1041A	TGGGGTCGGAGGCGAAGTAGTGGG	ATCCCGGAATACCCCGGCACAATTCACCACC
ChlH	S1103A	TTGACGTTGGAAGCATAGGAACCGGAGGCG	AATTGTTCCGGGGCATTCCGGGATCTG
ChlH	K1129A	TTGAAAGCGAAGGACGCGCGTTTGAGGTACATTTCC	ATGTACCTCAAACGCGCGTCCTTCGCTTTCAACTCC
ChlH	S1178A	TTGGTGGGGTCGGCGTCGAAGTAGTGG	CACTACTTCGACGCCGACCCCACCAAG