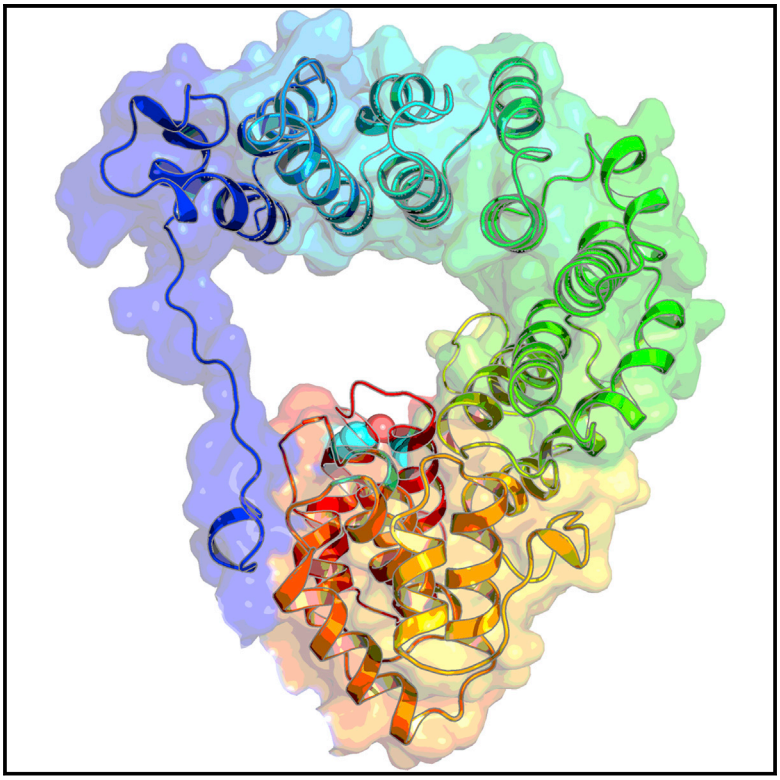
Structure

Article

Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase

Graphical abstract



Authors

Indika Kumarapperuma, Kes Lynn Joseph, Cong Wang, ..., Frédéric Partensky, Wendy M. Schluchter, Xiaojing Yang

Correspondence

xiaojing@uic.edu (X.Y.), wschluch@uno.edu (W.M.S.)

In brief

Chromophore attachment to the light-harvesting apparatus represents one of the most important post-translational modifications in photosynthetic cyanobacteria. By studying the crystal structures of enzymes catalyzing the chromophore attachment reaction, Kumarapperuma et al. have advanced our mechanistic understanding of how these enzymes perform and diversify their functions at the molecular level.

Highlights

- Crystal structure of an E/F-type bilin lyase reveals a "question-mark" architecture
- Structural and mutational studies establish active site for bilin ligation
- A tyrosine-mediated reaction scheme is proposed for bilin lyases
- Stereoselectivity plays a critical role in conferring the isomerase activity

Kumarapperuma et al., 2022, Structure *30*, 1–11
April 7, 2022 © 2022 Elsevier Ltd.
https://doi.org/10.1016/j.str.2022.01.007

CellPress

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

Structure

CellPress

Article

Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase

Indika Kumarapperuma,[1,5] Kes Lynn Joseph,[2,5] Cong Wang,[1] Linta M. Biju,[1] Irin P. Tom,[1] Kourtney D. Weaver,[2] Théophile Grébert,[3] Frédéric Partensky,[3] Wendy M. Schluchter,[2,*] and Xiaojing Yang[1,4,6,*]

[1]Department of Chemistry, University of Illinois Chicago, Chicago, IL 60607, USA
[2]Department of Biological Sciences, University of New Orleans, New Orleans, LA 70148, USA
[3]Ecology of Marine Plankton (ECOMAP) Team, Station Biologique, Sorbonne Université, CNRS, 29680 Roscoff, France
[4]Department of Ophthalmology and Vision Sciences, University of Illinois Chicago, Chicago, IL 60607, USA
[5]These authors contributed equally
[6]Lead contact
*Correspondence: xiaojing@uic.edu (X.Y.), wschluch@uno.edu (W.M.S.)
https://doi.org/10.1016/j.str.2022.01.007

SUMMARY

Chromophore attachment of the light-harvesting apparatus represents one of the most important post-translational modifications in photosynthetic cyanobacteria. Extensive pigment diversity of cyanobacteria critically depends on bilin lyases that covalently attach chemically distinct chromophores to phycobiliproteins. However, how bilin lyases catalyze bilin ligation reactions and how some lyases acquire additional isomerase abilities remain elusive at the molecular level. Here, we report the crystal structure of a representative bilin lyase-isomerase MpeQ. This structure has revealed a "question-mark" protein architecture that unambiguously establishes the active site conserved among the E/F-type bilin lyases. Based on structural, mutational, and modeling data, we demonstrate that stereoselectivity of the active site plays a critical role in conferring the isomerase activity of MpeQ. We further advance a tyrosine-mediated reaction scheme unifying different types of bilin lyases. These results suggest that lyases and isomerase actions of bilin lyases arise from two coupled molecular events of distinct origin.

INTRODUCTION

A critical step in the biogenesis of the light-harvesting complexes in photosynthetic cyanobacteria involves post-translational modifications that covalently attach bilin pigments to specific cysteine residues in phycobiliproteins (Schluchter et al., 2010). Such reactions are catalyzed by an important family of enzymes called *bilin lyases.* Phycobiliprotein α and β subunits incorporated with 1–3 bilins each are assembled as doughnut-shaped trimers or hexamers in megadalton phycobilisomes (PBSs) where the rods project out from the core to allow light harvesting and directional energy transfer to the reaction centers of photosystems (Glazer, 1989) (Figure S1A). Bilin lyases are directly responsible for the extensive pigment diversity of the PBSs and the ubiquity of cyanobacteria on Earth by enabling them to adapt and thrive in different light niches (Flombaum et al., 2013; Grébert et al., 2018; Sanfilippo et al., 2019a).

Among three phylogenetically distinct types of bilin lyases (Bretaudeau et al., 2012; Fairchild et al., 1992; Scheer and Zhao, 2008; Schluchter et al., 2010; Shen et al., 2006, 2008; Zhou et al., 1992), the S/U and T types have β-barrel folds, while the E/F type adopts an all-helical structure (Kronfel et al., 2013; Overkamp et al., 2014; Zhao et al., 2017; Zhou et al., 1992, 2014). These enzymes differ in their protein scaffolds and confer different substrate specificities for pigments, acceptor proteins, and target residues. In essence, they all catalyze the same chemical reaction, resulting in the formation of a thioether linkage between the sulfhydryl group of a cysteine residue in phycobiliprotein and the C3^1 atom in the A-ring of a bilin pigment (Scheer and Zhao, 2008; Schirmer et al., 1987) (Figure S1B). It has been proposed that this reaction involves a transient association between the bilin pigment and lyase, mediated by a nucleophilic residue such as histidine or cysteine from the lyase via the C10 atom (Stumpe et al., 1993; Tu et al., 2009; Zhao et al., 2000). However, the chemical identity of catalytic residues and their role in the enzymatic action remain elusive (Gasper et al., 2017; Overkamp et al., 2014; Scheer and Zhao, 2008; Tu et al., 2009; Zhao et al., 2000, 2017; Zhou et al., 2014). For E/F-type bilin lyases, only one structure has been solved for heterodimeric CpcE/F (Zhao et al., 2017).

To understand the molecular mechanism underlying the bilin ligation reaction, we investigate a single-chain, E/F-type bilin lyase MpeQ that incorporates phycourobilin (PUB) at the Cys83 site of the phycoerythrin-II α-subunit, a phycobiliprotein encoded by the *mpeBA* operon in marine *Synechococcus* (Grébert et al., 2021). The newly identified MpeQ is involved in the type IV chromatic acclimation (CA4) process, a widespread phenomenon in which marine *Synechococcus* varies the molar ratio between the blue-light-absorbing PUB and the green-light-absorbing phycoerythrobilin (PEB) in phycobiliproteins depending on the ambient light color (Everroad et al., 2006; Grébert

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

Table 1. Crystallography data collection and refinement statistics

Structure	MpeQ SeMet	MpeQ SeMet
Space group	C2	$C222_1$
Cell parameters		
a, *b*, *c* (Å)	83.84, 173.25, 114.42	136.67, 173.05, 113.7
α, β, γ (°)	90, 126.08, 90	90, 90, 90
Residues	1–398	1–398
Chain	A, B	A, B
Water		
Diffraction data		
X-ray source	21-ID-F, APS	21-ID-F, APS
Methods	monochromatic	monochromatic
	100 K	100 K
Wavelength (Å)	0.97872	0.97872
Resolution (Å)	56-2.5 (2.54–2.50)	50-2.90 (2.95–2.90)
R_{merge}	0.101 (0.605)	0.082 (0.761)
Completeness		
All (%)	99.5 (99.7)	99.8 (95.8)
Anomalous (%)	97.2 (84.7)	99.8 (95.5)
Redundancy (%)	3.7 (3.5)	3.9 (1.6)
$I/\sigma(I)$ (%)	7.4 (1.3)	4.7 (1.0)
Refinement		
Resolution (Å)	20-2.50 (2.79–2.50)	39.3-2.95 (3.01–2.95)
R	0.239 (0.279)	0.242 (0.361)
R_{free}	0.284 (0.329)	0.298 (0.303)
RMSD		
Bond length (Å)	0.009	0.008
Bond angle (°)	1.013	1.248
Average *B* (Å^2)	28.63	98.9
Ramachandran		
Favored (%)	96.6	96.1
Allowed (%)	2.0	3.3
Disallowed (%)	0.4	0.55
PDB entry	7MC4	7MCH

et al., 2018, 2021; Palenik, 2001; Sanfilippo et al., 2019a; Shukla et al., 2012). The CA4 phenomenon is attributed to light-regulated expression of a PEB lyase or lyase-isomerase encoded in a small CA4-conferring genomic island, which competes with a constitutively expressed homolog (lyase-isomerase or lyase) encoded in the large genomic region involved in the biosynthesis of phycobilisome rods (Grébert et al., 2021; Mahmoud et al., 2017; Sanfilippo et al., 2016, 2019a; Shukla et al., 2012). Compared with PEB lyases, bilin lyase-isomerases such as MpeQ have acquired an additional ability to isomerize PEB to PUB during the ligation reaction (Blot et al., 2009; Grébert et al., 2021; Shukla et al., 2012) (Figure S1C). However, the molecular basis for this isomerase action is not known. MpeQ thus offers a superb model system not only for dissecting the molecular mechanism of bilin lyases but also for addressing how these highly homologous enzymes attach different pigments to the same site of phycobiliproteins using the same substrate (Grébert et al., 2021; Humily et al., 2013; Sanfilippo et al., 2019b).

This work reports the crystal structure of MpeQ, a representative bilin lyase-isomerase from the E/F family. This structure has revealed a hitherto unknown "question-mark" architecture and active site geometry that call for a revision of the current model for E/F-type bilin lyases (Zhao et al., 2017). Based on structural analyses, site-directed mutagenesis, and enzyme assays, we identify key residues responsible for the lyase and isomerase activities. We propose a general tyrosine-mediated reaction scheme for bilin lyases in the E/F family and beyond. We further advance a model for the enzyme-substrate complex that elucidates how MpeQ accommodates both the bilin and protein substrates to catalyze the ligation reaction while employing stereoselectivity to confer the isomerase activity. These findings provide a structural framework and mechanistic insights into one of the most important enzymatic reactions in photosynthesis.

RESULTS

MpeQ adopts a question-mark architecture

We have determined the crystal structure of MpeQ at 2.5 Å resolution by the single-wavelength anomalous diffraction (SAD) method using Se-methionine-derivatized crystals (Table 1). In addition to the N-terminal histidine tag, all 398 residues of MpeQ have been accounted for in the electron density map with two protein molecules in an asymmetric unit. The α-solenoid structure of MpeQ adopts a question-mark scaffold, in which 23 α-helices are supercoiled in a right-handed manner (Figures 1A and 1D). The N-terminal (aa 14–255) and C-terminal (aa 274–398) domains connected by an extended linker are denoted as the E and F domains, respectively, corresponding to the CpcE and CpcF subunits of a prototypical CpcE/F bilin lyase from *Nostoc sp.* PCC 7120 (Zhao et al., 2017). However, the overall architecture of MpeQ starkly contrasts with the crescent-shaped structure proposed for CpcE/F (Zhao et al., 2017) (Figures 2A and 2B). Interestingly, but not surprisingly, a revised CpcE/F model based on an alternative arrangement of the E and F subunits presents a heterodimeric structure highly comparable to that of MpeQ (Figures 1B and 2). Remarkably, four out of five strictly conserved residues (Tyr76, Pro108, Arg112, and Trp151; residue numbering according toCpcF) between heterodimeric CpcE/F and single-chain E/F lyases such as MpeQ line up in the tight turns between helices facing an enclosed cavity in the question-mark architecture, in contrast to the published CpcE/F structure. This suggests that the MpeQ structure, but not the reported structure of CpcE/F (Zhao et al., 2017), represents the biologically relevant protein framework for bilin lyases in the E/F family.

Active site and key residues responsible for lyase activity

The E and F domains of MpeQ are roughly perpendicular to each other in terms of the overall helical direction (Figure 1C). The E domain constitutes the "stroke" part of the question mark, while the F domain corresponds to the "dot" part (Figures 1A and 1D). Besides the linker, the E and F domains are bridged by a long N-terminal extension (aa 1–13) enclosing a large interior chamber (Figures 1A and 3). In the crystal lattice, two molecules of MpeQ dimerize via their N-terminal extensions tethered as two

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

CellPress

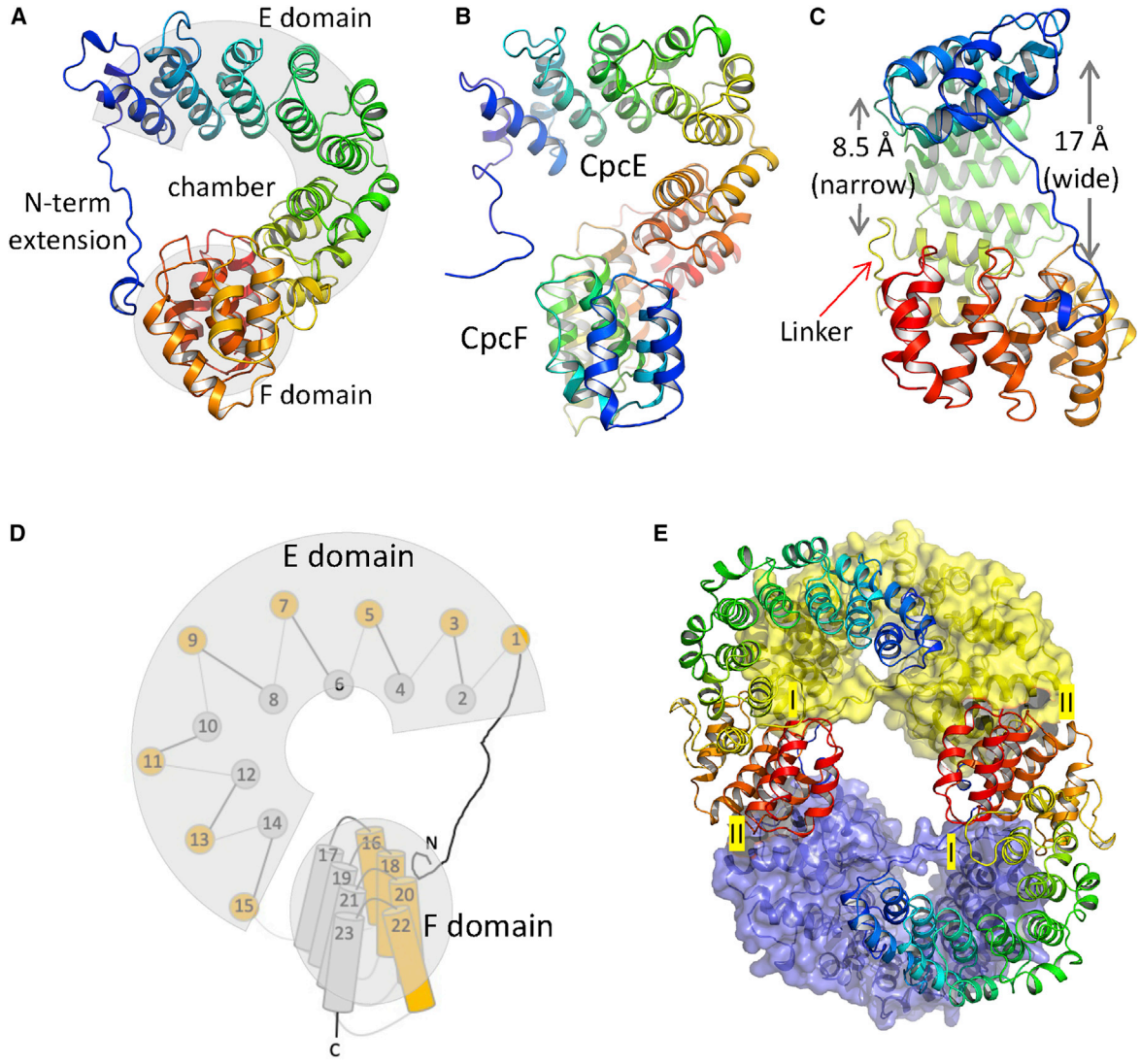


Figure 1. Protein architecture, topology, and crystal packing of MpeQ

(A) Ribbon diagram of MpeQ shows an α-solenoid "question-mark"-shaped structure with distinct E (aa 14–255) and F (aa 274–398) domains. A long N-terminal extension loops back to the F domain as a gate to the large interior chamber. The structure is depicted in rainbow colors from the N-terminus (blue) to C-terminus (red).

(B) The revised CpcE/F heterodimer (PDB: 5N3U) shows a protein architecture similar to that of MpeQ.

(C) A side view of MpeQ relative to (A) shows the wide and narrow openings of the chamber at opposite sides of MpeQ. The openings are about 8.5 Å on the narrow side (between Lys81 and Lys238) and 17 Å on the wide side (between Arg285 and Ile132), respectively.

(D) Topological diagram of the MpeQ structure is shown in a perspective to facilitate comparisons with CpcE/F (Zhao et al., 2017). A total of 23 helices shown in circles and rods highlight roughly perpendicular directions of helices in the E and F domains (shaded areas).

(E) Crystal packing of MpeQ shows two distinct intermolecular interfaces. Interface I (buried surface area of 2,426 Å^2) is much larger that interface II (buried surface area of 400 Å^2).

antiparallel β-strands traversing each other's chamber (Figure 1E). The enclosed chamber is open on both sides, although one side is notably wider than the other (Figure 1C). Arg/Lys residues pointing toward the chamber interior form large, positively charged surface patches on the ceiling and wall areas (Figures 3A and 4A). Along a shallow cleft between the E and F domains, some highly conserved residues, as well as those hallmark residues that distinguish the lyases and lyase-isomerases, are clustered on the chamber floor, constituting a putative active site (Figures 4B and S2). This site, marked by Tyr318, Lys353, and Glu285, is exposed to the wide side of the chamber (Figure 3B). In a crystal lattice, Tyr318 and Lys353 engage close interactions with the E′ domain of a neighboring molecule via an interface with a buried surface area of 2,426 Å^2 (Krissinel and Henrick, 2007) (Figure S3A). In solution, however, MpeQ is monomeric according to the elution profile of size exclusion chromatography (Figure S3B), which makes this site fully accessible for substrate binding.

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

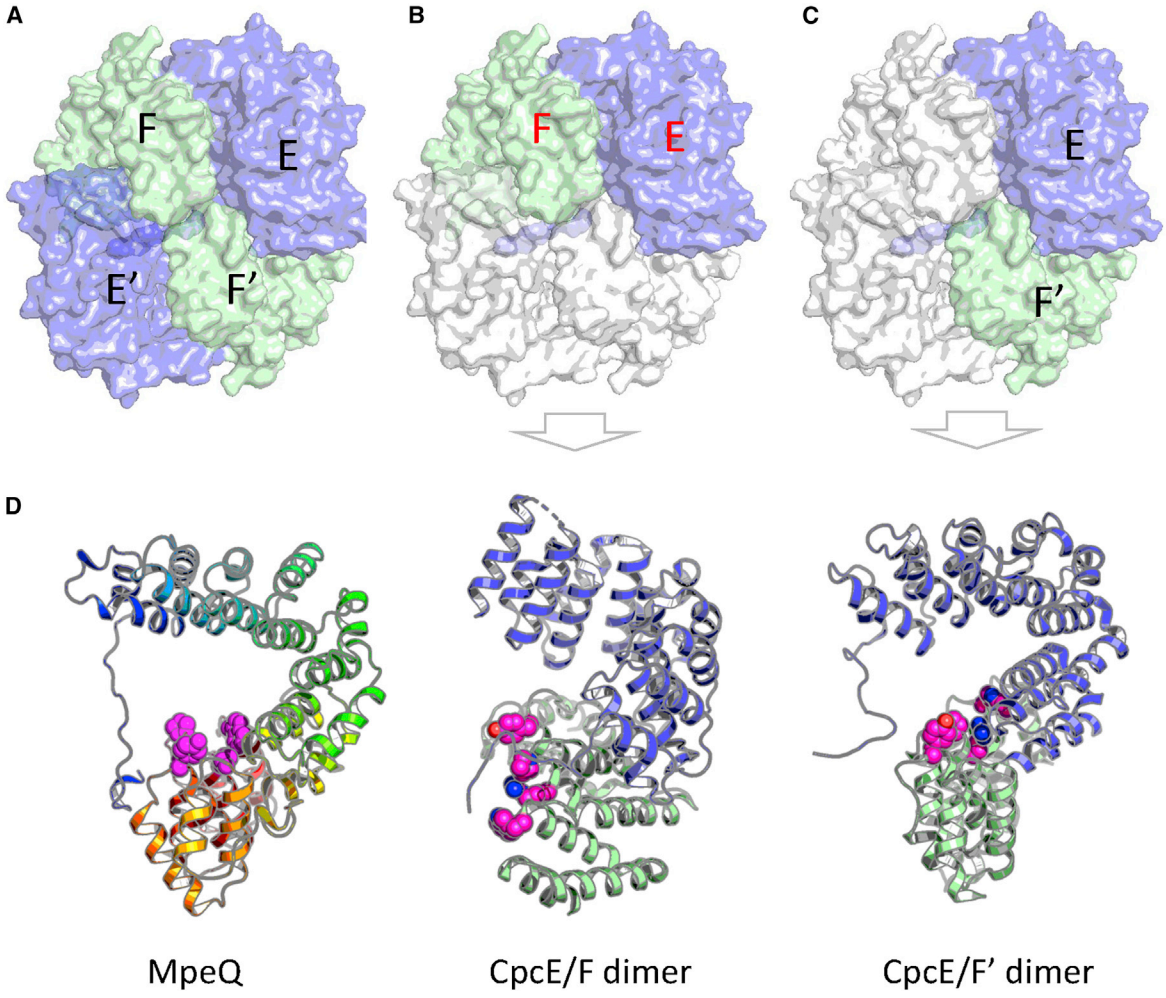


Figure 2. Crystal packing and dimer assemblies of CpcE/F

(A) Crystal packing of the CpcE/F structure shows two possible dimer assemblies of the E (blue) and F (green) domains (namely, the E/F assembly or E/F′ assembly). In crystallography, E and F are equivalent to E′ and F′, respectively.

(B) The CpcE/F assembly proposed by Zhao et al. (2017) (top) shows a crescent-shaped protein architecture (bottom).

(C) The revised CpcE/F′ assembly reveals a question-mark-shaped protein architecture (bottom).

(D) Comparisons between single-chain MpeQ (in rainbow) and two CpcE/F assemblies (E in blue; F in green). Magenta spheres mark the conserved residues (Tyr318, Pro349, Lys353, and Trp383: residue numbering according to MpeQ; Tyr76, Pro108, Arg112, and Trp151: residue numbering according to CpcF). See also Figure S5.

To confirm the active site, we made several single mutations in this cluster (Figures 4A and 4B) and examined how they affect the lyase and/or isomerase activities in MpeQ. By co-expressing MpeQ in *Escherichia coli* with two other plasmids that produce the MpeA and PEB substrates, we measured the normalized enzyme activities by detecting the chromophorylated products MpeA-PUB or MpeA-PEB using Zn-fluorescence (Berkelman and Lagarias, 1986) and/or absorption spectra of the purified MpeA (Figures 4C–4E and 5). The lyase activity was completely abolished in the single mutants of Y318F, Y318A, and K353A, with no detectable chromophorylation of MpeA (Figure 4C; Table S1). The E285A mutant showed significantly reduced lyase activity, while the S222A, S224P, and W383F variants retained the lyase-isomerase activities with MpeA-PUB produced at varying levels lower than the wild type (WT) (Figures 4C and 5). These mutational data corroborate our structural findings, supporting the idea that Tyr318 and Lys353 play essential roles in catalysis and/or substrate binding (Figure 3B).

We also made single mutants on several Arg residues lining the chamber interior (Figure 4A). When Arg71, Arg79, Arg135, or Arg143 was individually mutated to Asp, no MpeA chromophorylation was detected (Figure 5C). Similar results were observed for the R228A and R198A variants (Figures 4C and S2; Table S1). The R71A, R79A, R228D, and R198K variants showed low-level production of MpeA-PUB (Figure 5; Table S1). These results suggest that although Arg71, Arg79, Arg198, and Arg228 are not directly involved in catalysis, they are likely to act as protein anchors that help stabilize the bilin pigment via interactions with the negatively charged propionates.

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

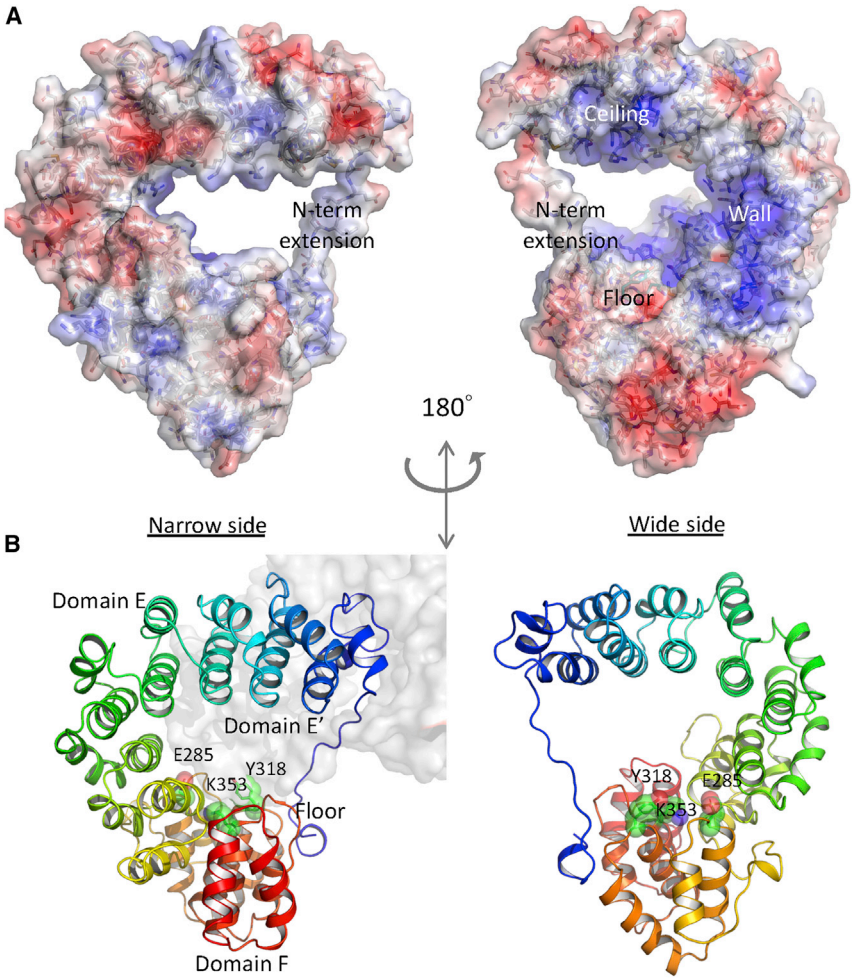


Figure 3. The active site chamber of MpeQ

(A) The electrostatic surface viewed from the narrow (left) and wide (right) sides. The chamber interior is largely positively charged where Arg/Lys residues are clustered in the ceiling, wall, and floor regions of the chamber.

(B) The catalytic triad on the floor consists of Tyr318, Lys353, and Glu285 (green spheres) located in the loops between helices in the F domain.

Steric factors underlying isomerase activity

To determine the molecular basis of isomerase activity, we analyzed a collection of protein sequences of highly homologous E/F-type lyases and their counterpart lyase-isomerases. Specifically, sequence comparisons of the MpeW and MpeY lyases with the MpeQ and MpeZ lyase-isomerases retrieved from 106 marine *Synechococcus* strains or single-amplified genomes (Data S1) highlighted nine residues swapped between the lyases and lyase-isomerases (Figure S2; Data S2). To examine whether these swapped residues contribute to the gain of the isomerase function, we individually substituted five such residues at the active site of MpeQ with their counterparts in the PEB lyase MpeW (Figure S4). All five single mutants (A100T, S224P, V319G, T320A, and Y323Q) largely retained their lyase activity (Figures 5 and S4). Only the V319G variant produced an additional absorption peak near 550 nm, indicating the formation of MpeA-PEB (Figure S4B). In other words, the Val→Gly substitution led to partial loss of the isomerase activity in MpeQ (Figures 4D and S4B).

Val319 is located right next to the catalytic Tyr318. We systematically explored possible steric effects on lyase-isomerase activity by replacing Val319 with Gly, Ala, Leu, and Phe individually (Figures 4C and 4D). With increasingly bulkier side chains at the position of Val319, these mutants exhibited a striking color-tuning effect on the MpeA product (Figure 5A). The V319G variant with the smallest side chain showed a reddish color, indicating a low PUB:PEB ratio in MpeA. The V319A variant with a slightly larger side chain is orange, suggesting a higher PUB:PEB ratio. With valine or leucine present at position 319, the yellow sample suggests the exclusive formation of MpeA-PUB. Larger side chains at this position appear to interfere with the bilin attachment. As such, the lyase activity of the V319L variant was reduced by about 50%, while the V319F variant barely showed any detectable chromophorylated MpeA (Figures 4C, 4D, and 5).

It is not surprising that the single swap mutations (A100T, S224P, T320A, and Y323Q) did not directly affect the isomerase function, as these sites are farther from the catalytic Tyr318 (Figures 4B and S4B). However, when these substitutions were made in combination with V319G, they significantly altered the PUB:PEB ratio in the MpeA products (Figure 4E). The single mutant V319G exhibited a PUB:PEB ratio of 10:1, while the double mutant V319G/T320A lowered the PUB:PEB ratio to 3:1. This ratio was further reduced to 1.5:1 in the triple mutant V319G/T320A/Y323Q. Two quintuple mutants (MpeQ5.1: A100T/V319G/T320A/Y323Q/T352A and MpeQ5.2: S224P/V319G/T320A/Y323Q/T352A) evidently switched to make more MpeA-PEB than MpeA-PUB, resulting in PUB:PEB ratios of 1:1.5 and 1:3, respectively (Figure 4E). It is noteworthy that higher PUB:PEB ratios coincide with bulkier side chains at these swap sites (Figure 4A), suggesting that the formation of MpeA-PEB requires a larger reaction volume at the active site. To this end, exclusive attachment of PEB to MpeA-Cys83 by MpeW arises from collective steric effects from multiple swap sites beyond Val319, as exemplified by the sequence differences between MpeQ and its counterpart lyase MpeW (Figures S2 and S5).

Tripartite model and reaction mechanism

Identification of the active site allowed us to explore the modes of substrate binding to MpeQ. We first examined the structural flexibility of both MpeQ and MpeA using molecular dynamics (MD) simulations (Figure S6). A 72 ns MD simulation on MpeQ revealed two distinct dispositions of the long N-terminal extension (~aa 1–15) (Figure S6A). In a "latch-on" position similar to that in

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

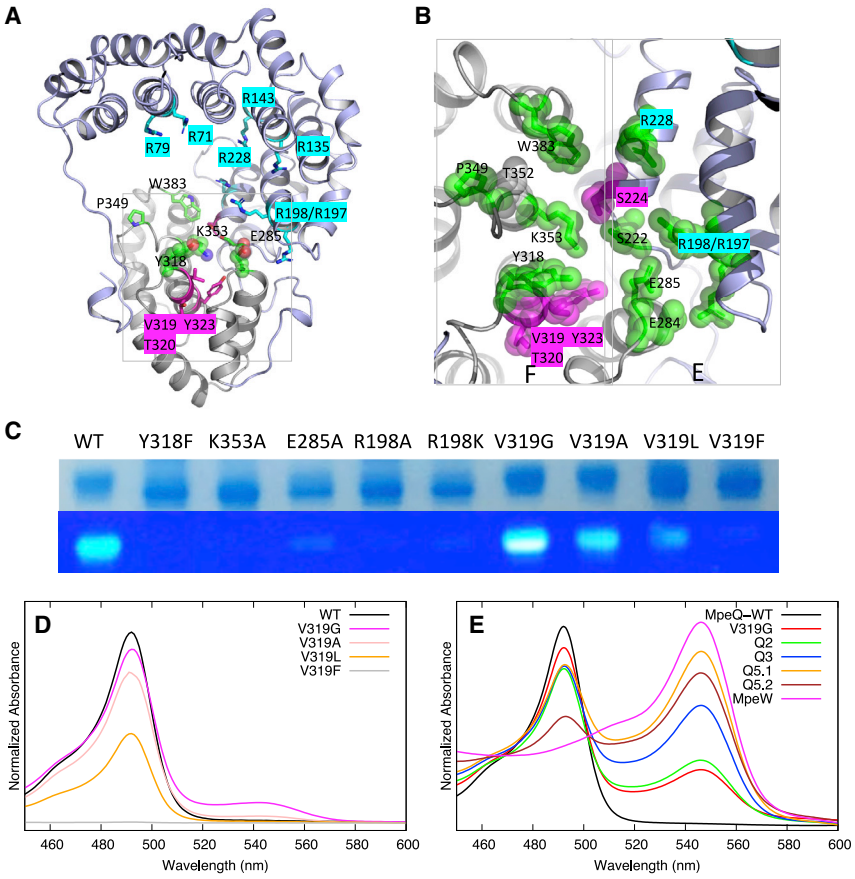


Figure 4. Site-directed mutagenesis in the active site

(A) Sites of mutation are highlighted in stick models: arginines (cyan), swap residues (magenta), and catalytic residues (green spheres).

(B) Conserved (green) and swap (magenta) residues form a surface patch at the chamber floor at the E/F interface.

(C) Coomassie blue staining (top) and Zinc fluorescence (bottom) of the same SDS-PAGE gel of MpeA obtained via co-expression with MpeQ.

(D) Normalized absorption spectra of MpeA obtained from the MpeQ mutants at Val319.

(E) MpeQ variants carrying one or more swap mutations show varying PEB:PUB ratios. Q2: V319G/T320A; Q3: V319G/T320A/Y323Q; Q5.1: A100T/V319G/T320A/Y323Q/T352A and Q5.2: S224P/V319G/T320A/Y323Q/T352A. MpeW is the PEB lyase counterpart of MpeQ.

the crystal structure, the N-terminal extension encloses the active site chamber, while the "latch-off" position allows better accessibility to the active site. The root-mean-square deviation (RMSD) distance matrix derived from 100 simulated structures on the trajectory reveals three rigid bodies in the α-solenoid structure (aa 30–97 and 117–237 in the E domain; aa 250–398 in the F domain), which roughly correspond to the ceiling, wall, and floor sections of the chamber, respectively (Figures S6B and 3B). The MD trajectory shows close coupling between the wall and floor sections, while the ceiling section exhibits higher mobility, likely due to its proximity to the flexible N-terminal extension. We also performed the MD simulations on an apo-MpeA model (Swiss model) (Biasini et al., 2014), which reveals two distinct clusters of conformations in the caplike loop (aa 63–81), shielding MpeA-Cys83 from the molecular surface (Figures S6C and S1B). The "cap-on" conformation resembles a well-folded α-subunit, while the "cap-off" structure exposes MpeA-Cys83 for docking at the active site of MpeQ.

Based on surface complementarity, protein flexibility, and active site geometry, we present a tripartite model in which MpeQ is complexed with the PEB and MpeA substrates (Figure 6). First, the active site in monomeric MpeQ is fully accessible for substrate binding (Figure 6A). Apo-MpeA is expected to approach the active site chamber from the wide side. This docking is likely facilitated by induced fit involving both MpeA and MpeQ (Figures 6A and S6). Specifically, the N-terminal extension of MpeQ would adopt a latch-off position that allows proper positioning of MpeA-Cys83 at the active site (Figure 6B). With the bilin binding pocket of MpeA open to the chamber interior, the active site of MpeQ is accessible to the PEB chromophore from the narrow side (Figure 6C). This model features PEB in its final binding cleft of MpeA (Figures 6B and 6C), with its A-ring sandwiched between MpeQ-Tyr318 and MpeA-Cys83.

At the active site, the A-ring assumes an extended *anti*-conformation as in the final product. This disposition presents the C3=C3^1 double bond, the target site for ligation, to the hydroxyl group of Tyr318, while its lactam group is positioned to engage direct interactions with Lys353/Glu285 (Figures 6D; and S7A). This active site geometry fully agrees with the mutagenesis data (Figure 4C), supporting the idea that Tyr318 is critical for catalyzing the ligation reaction, while Lys353/Glu285 are important for conferring a specific orientation of the A-ring. In addition, this model places Val319 right next to the *R*-configured C2 atom, which perfectly explains the remarkable steric effects of Val319 on the lyase and isomerase activities (Figure 4D).

This tripartite model clearly points to distinct chemical origins for the lyase and isomerase actions of MpeQ. To address whether the PEB → PUB isomerization occurs before or after bilin ligation to MpeA-Cys83, we performed a simple experiment by coexpressing MpeQ with a PEB-producing plasmid and the MpeA-C83A variant. We reason that since MpeA-C83A is unable to covalently attach the bilin pigment, the ligation reaction cannot complete, thereby allowing us to detect the isomerization product PUB if it is formed in the absence of ligation. We measured absorption spectra from the load, flowthrough, and eluted fractions of the affinity chromatography column that was used to separate bulk proteins and free pigments from the His-tagged MpeQ and MpeA (Figure S8A). Both the load and flowthrough samples showed the absorption band around 560 nm, supporting the idea that the PEB substrate was

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

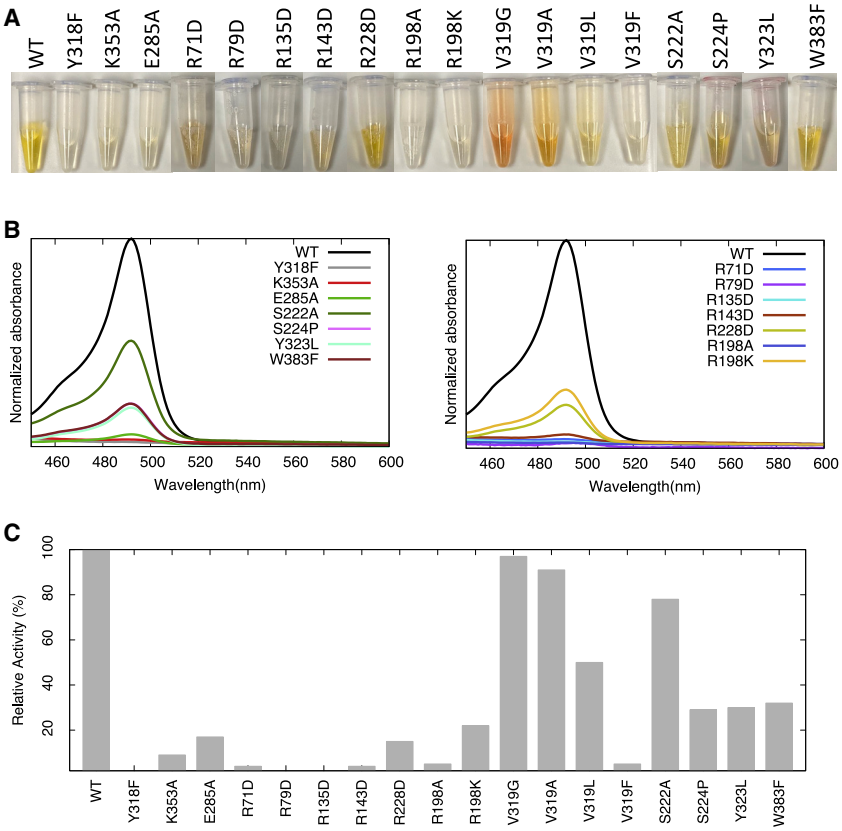


Figure 5. Lyase and lyase-isomerase activities of the MpeQ variants in an *E. coli* coexpression system

(A) The distinct sample color of MpeA co-expressed with MpeQ serves as a good indicator of the lyase and lyase-isomerase activities of MpeQ. The yellow color of MpeA-PUB is attributed to the lyase-isomerase activity, while the red/orange color of MpeA-PEB is a result of additional lyase activity as shown for V319G.

(B) Normalized absorption spectra of MpeA co-expressed with various single mutants of MpeQ.

(C) Relative activities of the MpeQ variants. Absorbance at 495 nm normalized by relative band intensities of MpeA (Figure 4C) is used to quantify the lyase activity. All relative activities are obtained in reference to that of the wild type (WT) MpeQ (set to 100%).

sufficiently available in this co-expression system. While PUB was barely detectable in the bulk samples of load and flow-through, the eluted fraction exhibits a small yet clear peak at ~495 nm (Figure S8A), suggesting that PUB is formed in the absence of ligation to MpeA-Cys83 and remains associated with MpeQ and/or MpeA.

Taken together, we advance a reaction scheme to elucidate a tyrosine-mediated bilin ligation and isomerization catalyzed by MpeQ. In essence, the net reaction catalyzed by a bilin lyase is a nucleophilic addition of the sulfhydryl (-SH) group of cysteine in the phycobiliprotein substrate to the C3=C3^1 double bond of the bilin substrate (Figure 7). In the final product of MpeQ, a thioether bond is formed between the C3^1 atom and the sulfur atom of MpeA-Cys83, while the C3=C3^1 double bond is reduced. Such a nucleophilic addition to alkene requires a catalytic group to polarize the C=C double bond. Our structural and mutagenesis data strongly suggests that the hydroxyl group of MpeQ-Tyr318 serves this role. We propose that in the first step of the reaction, Tyr318 activates the bilin substrate, resulting in a polarized C3=C3^1 double bond (Figure 7). This bilin activation step is critical, as no lyase activity was detected in the Y318F or Y318A mutants, when this hydroxyl group is absent, or in V319F, where the catalytic hydroxyl group cannot get close enough to the C3 atom (Figure 4). In MpeW, where the active site poses no major steric hinderance, the bilin activation proceeds to the nucleophilic addition reaction, giving rise to the MpeA-PEB product. In MpeQ, on the other hand, the MpeA-PEB formation is not allowed due to incompatibility between the active site residue (Val319) and the chiral centers at the C3 and/or C2 positions in PEB. The addition reaction cannot take place unless this conflict is resolved in a rearrangement reaction where PEB→PUB isomerization effectively shifts the chiral centers from the C2/C3 to C4/C5 positions (Figures 6 and 7). Selected by evolution, the side chain of Val319 is small enough to permit the catalytic action of Tyr318, yet bulky enough to exclude the PEB attachment. Similar to the retinoid isomerase RPE65 (Kiser et al., 2015), MpeQ exemplifies a general molecular strategy to confer isomerase activity via stereoselectivity of the active site geometry (Figure 7).

DISCUSSION

Prior to this study, no crystal structure had been reported for any bilin lyase-isomerase, and CpcE/F (a PCB lyase from *Nostoc sp.* PCC 7120) was the only structure available for the E/F-type bilin lyases (Zhao et al., 2017). However, the interface between CpcE and CpcF proposed by Zhao et al. is not present in the structure of single-chain MpeQ (Figures 1E and 2B) (Zhao et al., 2017). In fact, no consistent pattern or interface other than the question-mark arrangement is found between the crystal structures of MpeQ and CpcE/F (Figure 2D). Therefore, it is unlikely that the crescent-shaped architecture (PDB: 5N3U) represents the biological assembly or functional unit of CpcE/F despite its high resolution. Because the crystal lattice of CpcE/F presents two crystallography-equivalent yet distinct dimeric arrangements (Figure 2), it is susceptible to misinterpretation of the biologically relevant heterodimer. The single-chain MpeQ, on the other hand, is not prone to such ambiguity, especially when all residues have been accounted for in the electron density map. Given the similarities in protein sequence and secondary structure (Figures 1 and S5), we posit that the E/F family of bilin lyases, including the single-chain and heterodimeric enzymes, share a common structural framework represented by the question-mark architecture of MpeQ.

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

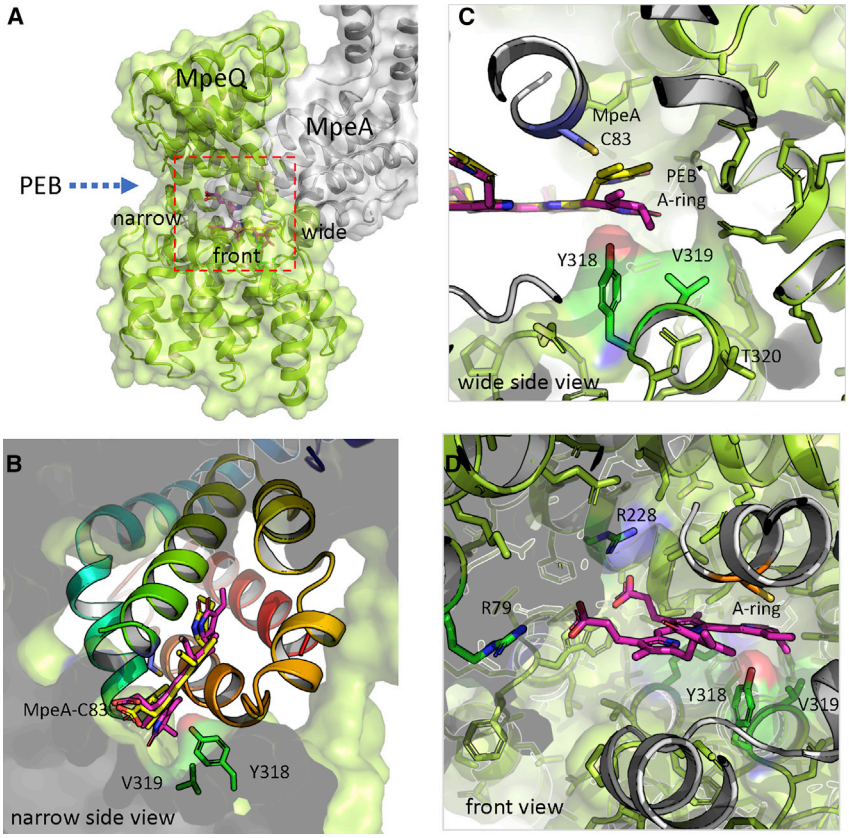


Figure 6. A docking model for substrates at the active site

(A) The active site of MpeQ (green) is accessible from both sides of the chamber. MpeA (gray) approaches the active site from the wide opening, while the PEB substrate enters from the narrow side (blue arrow).

(B) Viewed from the narrow side, the bound bilin (PEB in magenta; PUB in yellow) is docked at the final binding cleft in the MpeA structure (rainbow ribbon diagram).

(C) Viewed from the wide side, the catalytic Tyr318 is poised to activate the A-ring, which causes a steric clash between PEB and Val319. MpeQ resolves this conflict via PEB→PUB isomerization, which presents the A-ring to MpeA-C83 for ligation.

(D) In this front view, PEB is docked with its ring's A/B side lying in the cleft between the E and F domains, while the propionate groups are stabilized by Arg79 and Arg228.

We also note remarkable similarities in the active site geometry between MpeQ and other types of bilin lyases despite their distinct protein scaffolds (Figure S7A). The catalytic triad of MpeQ is found not only in the revised structure of CpcE/F but also in CpcT/CpeT (PDB: 4O4O, 4O4S, and 5HI8) and CpcS/CpeS (PDB: 3BDR and 4TQ2) (Gasper et al., 2017; Kronfel et al., 2013; Overkamp et al., 2014; Zhao et al., 2017; Zhou et al., 2014). Specifically, they all feature a conserved tyrosine at the rim of a large protein cavity where the phycobiliprotein and bilin substrates are supposed to bind. And CpcF-Tyr76 is perfectly aligned with MpeQ-Tyr318 (Figures S5 and S7A). Beyond its locality, this conserved active-site tyrosine seems to play a similar role. Just like MpeQ-Y318F, no bilin attachment was detected in the single mutants of CpcF-Y76A and CpcT-Y65F (Zhao et al., 2017; Zhou et al., 2014). We speculate that the proposed tyrosine-mediated reaction mechanism also applies to bilin lyases in the T and S/U clans (Figure 7).

Regarding the origin of the isomerase activity, we have considered two scenarios. First, the E/F lyase and isomerase activities result from two distinct enzymatic reactions involving different catalytic residues, and these swap residues simply play a role in tipping off the balance between these two reactions. Second, the isomerase activity is an add-on function associated with the lyase reaction catalyzed by Tyr318, which arises from collective steric effects from the swap residues near the active site. The first scenario implies that the catalytic residues responsible for the isomerase activity are near the C4=C5 double bond of PEB and that they are distinct from Tyr318. However, our current docking model shows that the only residues within a 5 Å radius of the C4=C5 double bond of PEB are MpeQ-Tyr318 (3.5 Å) and MpeA-Cys83 (4.4 Å) (Figure 6). In addition, it is unlikely that MpeA-Cys83 is directly involved in isomerization (Figure S8A). On the other hand, a bioinformatic study on isomerases (Martinez Cuesta et al., 2014; Martínez Cuesta et al., 2016) suggested that the isomerase and lyase activities are evolutionarily related due to their similarity in chemistry, which goes along with the second scenario. We propose that highly homologous MpeQ and MpeW share similar reaction mechanisms for their lyase actions, while the substrate transformation between the activation and ligation steps is directly influenced by the geometry and steric factors of the active site (Figure 7).

Our mutagenesis data shows that the steric factor at the position of Val319 directly influences the isomerase activities (Figures 4C, 4D, S8B and S8C). Consistently, among the 106 sequences of the MpeQ/W/Y/Z family retrieved from marine *Synechococcus* strains or single-amplified genomes (SAGs)(-Grébert et al., 2021; Sanfilippo et al., 2019a; Shukla et al., 2012), all enzymes with the isomerase function have a valine residue at the this position, whereas all lyases have a glycine (Figure S2; Data S2). This includes MpeZ, a lyase-isomerase shown to be involved in the CA4 strain RS9916 (Sanfilippo et al., 2019a; Shukla et al., 2012), which features Val next to the catalytic Tyr321, while its counterpart PEB lyase MpeY has Gly instead. However, the Val→Gly substitution alone is not sufficient (even in combination with some swap sites) to confer exclusive PEB attachment in MpeQ (Figure 4E). Evidently, a complete switch from MpeQ to a MpeW-like PEB lyase requires contributions from other swap sites, which collectively confer an active site that is spacious enough to accommodate the C2/C3 chiral centers of a non-planar A-ring during the ligation reaction (Figures 4E; and S2).

It is worth noting that the heterodimeric lyase-isomerase PecE/F lacks the tyrosine corresponding to Tyr318. Homology

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

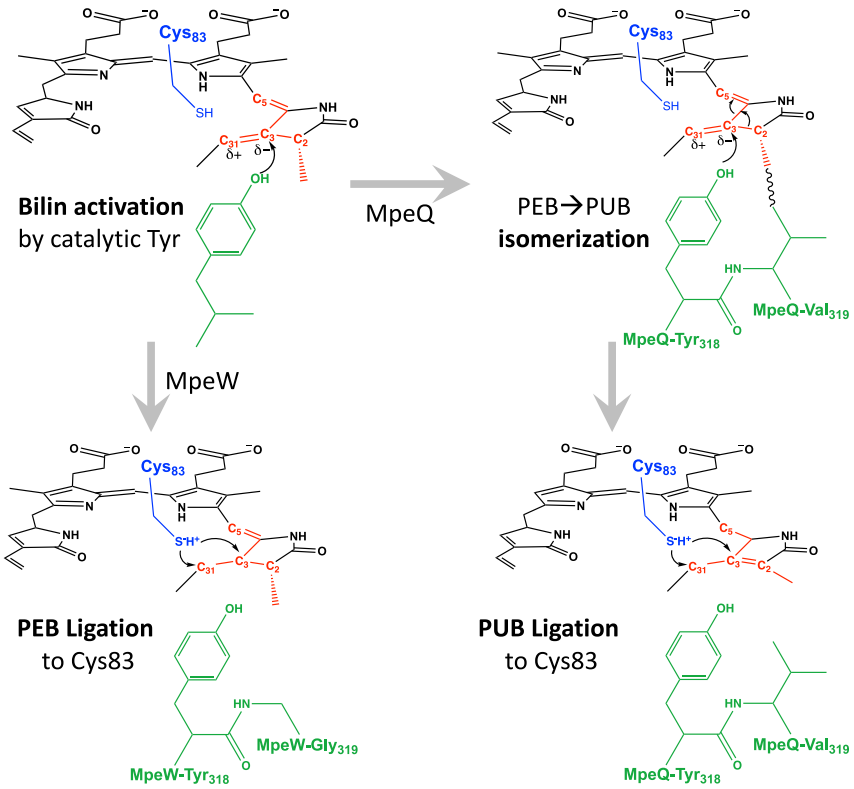


Figure 7. A proposed reaction scheme forbilin addition reaction catalyzed by bilin lyase MpeW and lyase-isomerase MpeQ

In the first step of the reaction, the bilin substrate PEB is activated by a catalytic tyrosine in both MpeW and MpeQ, resulting in a polarized C3=C3^1 double bond. In MpeW, Gly319 next to the catalytic Tyr318 is compatible with the non-planar A-ring conformation of PEB. The activated PEB is thus allowed to proceed to the nucleophilic addition reaction, forming a thioether bond between MpeA-Cys83 (blue) and the A-ring C3^1 atom. In MpeQ, the side chain of Val319 renders a steric conflict with the A-ring of PEB. The ligation reaction proceeds only when this conflict is resolved via the PEB→PUB isomerization, resulting in the MpeA-PUB product.

modeling of PecE/F places its signature HC motif right at the space occupied by MpeQ-Tyr318 (SwissModel) (Figure S7B). In addition, the single mutations in the HC motif abolished the lyase activity of PecE/F (Zhao et al., 2017). It is possible that PecE/F employs histidine instead of tyrosine as a catalytic residue for substrate activation (Tu et al., 2009) (Figure S7B). This exception lends further support to the active site geometry proposed for the E/F bilin lyases.

The active site geometry as illustrated in our tripartite model not only supports the respective roles of Tyr318 and Val319 but also has other implications. It is plausible that Lys353/Glu285 from the catalytic triad contribute to stabilizing the A-ring in a *5,anti*-conformation via hydrogen bond interactions with the lactam group (Figures S7A and 6D). This feature hints on the hitherto unknown structural basis for the *5,anti*-conformation of bilins exclusively found in phycobiliproteins, in contrast to the 5,*syn*-conformation observed in phytochromes (Wagner et al., 2005; Yang et al., 2007). In other words, the characteristic *5,anti*-conformation arises from substrate selectivity of the bilin lyases. As such, altered protein-chromophore interactions in the K353A and E285A mutants resulted in poor bilin attachment (Figure 4C). Furthermore, positively charged residues (e.g., Arg79/Arg228) are positioned close to the bilin propionates in the active site chamber (Figure 6D); such ionic interactions are reminiscent of the bilin binding modes observed in phytochromes and other bilin-binding proteins (Wagner et al., 2005; Yang et al., 2008). Not surprisingly, the R79D and R228D mutants adversely affected the lyase activities of MpeQ (Figure 5).

Needless to say, further studies are needed to validate the atomic details of the proposed reaction scheme, which are beyond the scope of this work. To this end, our attempts to directly introduce substrates to the MpeQ crystals by soaking have been unsuccessful because the current crystal form blocks any substrate binding to the active site (Figures 1E and S3A). MpeQ complexed with pigment alone is also deemed unfeasible for co-crystallization due to the large size difference between the active site chamber and PEB. We reason that both substrates (MpeA and PEB) are needed to assemble stable complexes for structural studies by crystallography or cryoelectron microscopy. To accumulate homogeneous species for static studies, the ligation reaction can be arrested by using either an inactive enzyme such as MpeQ-Y318F or a disabled substrate such as MpeA-C83A. For example, MpeQ-Y318F can be used to capture the preactivation state, while MpeA-C83A allows trapping of later preligation intermediates. It will be intriguing to see whether bilin activation mediated by Tyr318 involves any formation of a tyrosine radical and/or transient covalent intermediates.

Why do some E/F lyases such as CpcE/F employ two polypeptide chains while other homologous enzymes such as MpeQ achieve the same function using a single protein chain? Biswas et al. suggested that the CpcE/F family evolved after the CpcS/U family of lyases because some E/F members have the ability to remove bilins (Biswas et al., 2011). We speculate that E/F heterodimers have advantages in allowing mix-and-match of different E- and F-like subunits. It is noteworthy that the gene fusion events seem to occur more frequently in phycoerythrin-containing organisms. For example, RpcG (a PEB lyase-isomerase) is a fusion protein, with its N-terminal and C-terminal domains closely related to PecE and PecF, respectively (Blot et al., 2009), while CpeY (a PEB lyase for α-phycoerythrin I) appears to result from the fusion of CpcE- and CpcF-like domains (Biswas et al., 2011). These examples provide clues about how the E/F clan shifted from mainly heterodimers to single-chain enzymes, the latter conformation being the rule for all phycoerythrin-related lyases and lyase-isomerases.

Taken together, the results from this study have revealed distinct chemical origins for the lyase and isomerase actions of bilin lyase-isomerase in the E/F family. Our structural and mutagenesis data strongly support the finding that stereoselectivity of the

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

active site geometry plays a critical role in conferring the isomerase activity of MpeQ. We further propose a tyrosine-mediated reaction scheme shared among different types of bilin lyases. This work thus presents a structural and mechanistic framework that has advanced the fundamental understanding of critical biochemical reactions underlying the phycobilisome biogenesis.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Protein purification and site-directed mutagenesis
  - Crystallization, data collection and structure determination
  - Sequence analysis and identification of switch residues
  - Enzyme activity assays based on a co-expression system
  - Molecular dynamics (MD) simulations
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2022.01.007.

ACKNOWLEDGMENTS

We thank the staff of the Life Science Consortium Access Team (LS-CAT) at the Advanced Photon Source (APS) for support in X-ray diffraction data collection. Use of the LS-CAT Sector 21 is supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor under grant 085P1000817. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357. This work is supported by the National Science Foundation grant MCB 2017171 to W.M.S. and MCB 2017274 to X.Y., the National Institutes of Health grant R01EY024363 to X.Y., and the French National Agency for Research (ANR) program EFFICACY (ANR-19-CE02-0019) to F.P.

AUTHOR CONTRIBUTIONS

Conceptualization, X.Y. and W.M.S.; methodology, I.K., K.L.J., W.M.S., and X.Y.; investigation, I.K., K.L.J., L.M.B., I.P.T., K.D.W., C.W., W.M.S., and X.Y.; analysis, I.K., K.L.J., T.G., F.P., W.M.S., and X.Y.; writing – original draft, X.Y.; writing – review & editing, X.Y., W.M.S., T.G., and F.P.; funding acquisition, X.Y., W.M.S., and F.P.; resources, I.K. and K.L.J.; supervision, X.Y. and W.M.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science.

Received: September 10, 2021
Revised: December 1, 2021
Accepted: January 17, 2022
Published: February 10, 2022

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Berkelman, T.R., and Lagarias, J.C. (1986). Visualization of bilin-linked peptides and proteins in polyacrylamide gels. Anal. Biochem. *156*, 194–201.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. *42*, W252–W258.

Biswas, A., Boutaghou, M.N., Alvey, R.M., Kronfel, C.M., Cole, R.B., Bryant, D.A., and Schluchter, W.M. (2011). Characterization of the activities of the CpeY, CpeZ, and CpeS bilin lyases in phycoerythrin biosynthesis in Fremyella diplosiphon strain UTEX 481. J. Biol. Chem. *286*, 35509–35521.

Blot, N., Wu, X.-J., Thomas, J.-C., Zhang, J., Garczarek, L., Böhm, S., Tu, J.-M., Zhou, M., Plöscher, M., Eichacker, L., et al. (2009). Phycourobilin in Trichromatic phycocyanin from Oceanic cyanobacteria is formed post-translationally by a phycoerythrobilin lyase-isomerase. J. Biol. Chem. *284*, 9290–9298.

Bretaudeau, A., Coste, F., Humily, F., Garczarek, L., Le Corguillé, G., Six, C., Ratin, M., Collin, O., Schluchter, W.M., and Partensky, F. (2012). CyanoLyase: a database of phycobilin lyase sequences, motifs and functions. Nucleic Acids Res. *41*, D396–D401.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

Crooks, G.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

DeLano, W.L. (2002). Pymol: an open-source molecular graphics tool. CCP4 Newsl. Protein Crystallogr. 82–92.

Doublié, S. (1997). [29] Preparation of selenomethionyl proteins for phase determination. In Methods in Enzymology, C.W. Carter, Jr., ed. (Academic Press), pp. 523–530.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Everroad, C., Six, C., Partensky, F., Thomas, J.-C., Holtzendorff, J., and Wood, A.M. (2006). Biochemical bases of type IV chromatic adaptation in marine Synechococcus spp. J. Bacteriol. *188*, 3345–3356.

Fairchild, C.D., ZHAOt, J., ZHOUtl, J., Colson, S.E., BRYANTt, D.A., and Glazer, A.N. (1992). Phycocyanin a-subunit phycocyanobilin lyase. Proc. Natl. Acad. Sci. U S A *89*, 7017–7021.

Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincon, J., Zabala, L.L., Jiao, N., Karl, D.M., Li, W.K.W., Lomas, M.W., Veneziano, D., et al. (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. Proc. Natl. Acad. Sci. U S A *110*, 9824–9829.

Gasper, R., Schwach, J., Hartmann, J., Holtkamp, A., Wiethaus, J., Riedel, N., Hofmann, E., and Frankenberg-Dinkel, N. (2017). Distinct features of Cyanophage-encoded T-type phycobiliprotein lyase ΦCpeT: the role OF auxiliary metabolic genes. J. Biol. Chem. *292*, 3089–3098.

Glazer, A.N. (1989). Light guides. Directional energy transfer in a photosynthetic antenna. J. Biol. Chem. *264*, 1–4.

Grébert, T., Doré, H., Partensky, F., Farrant, G.K., Boss, E.S., Picheral, M., Guidi, L., Pesant, S., Scanlan, D.J., Wincker, P., et al. (2018). Light color acclimation is a key process in the global ocean distribution of *Synechococcus cyanobacteria*. Proc. Natl. Acad. Sci. U S A *115*, E2010–E2019.

Grébert, T., Nguyen, A.A., Pokhrel, S., Joseph, K.L., Ratin, M., Dufour, L., Chen, B., Haney, A.M., Karty, J.A., Trinidad, J.C., et al. (2021). Molecular bases of an alternative dual-enzyme system for light color acclimation of marine *Synechococcus* cyanobacteria. Proc. Natl. Acad. Sci. U S A *118*, e2019715118.

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., and MacKerell, A.D. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat. Methods *14*, 71–73.

Humily, F., Partensky, F., Six, C., Farrant, G.K., Ratin, M., Marie, D., and Garczarek, L. (2013). A gene island with two possible Configurations is involved in chromatic acclimation in marine Synechococcus. PLoS One *8*, e84459.

Hussain, H., and Chong, N.F.-M. (2016). Combined overlap extension PCR method for improved site directed mutagenesis. Biomed. Res. Int. *2016*, 1–7.

Kiser, P.D., Zhang, J., Badiee, M., Li, Q., Shi, W., Sui, X., Golczak, M., Tochtrop, G.P., and Palczewski, K. (2015). Catalytic mechanism of a retinoid isomerase essential for vertebrate vision. Nat. Chem. Biol. *11*, 409–415.

Kronfel, C.M., Kuzin, A.P., Forouhar, F., Biswas, A., Su, M., Lew, S., Seetharaman, J., Xiao, R., Everett, J.K., Ma, L.-C., et al. (2013). Structural and biochemical characterization of the bilin lyase CpcS from Thermosynechococcus elongatus. Biochemistry *52*, 8663–8676.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Kronfel, C.M., Biswas, A., Frick, J.P., Gutu, A., Blensdorf, T., Karty, J.A., Kehoe, D.M., and Schluchter, W.M. (2019). The roles of the chaperone-like protein CpeZ and the phycoerythrobilin lyase CpeY in phycoerythrin biogenesis. Biochim. Biophys. Acta *1860*, 549–561.

Mahmoud, R.M., Sanfilippo, J.E., Nguyen, A.A., Strnat, J.A., Partensky, F., Garczarek, L., Abo El Kassem, N., Kehoe, D.M., and Schluchter, W.M. (2017). Adaptation to blue light in marine Synechococcus requires MpeU, an enzyme with similarity to phycoerythrobilin lyase isomerases. Front. Microbiol. *8*, 243.

Martinez Cuesta, S., Furnham, N., Rahman, S.A., Sillitoe, I., and Thornton, J.M. (2014). The evolution of enzyme function in the isomerases. Curr. Opin. Struct. Biol. *26*, 121–130.

Martínez Cuesta, S., Rahman, S.A., and Thornton, J.M. (2016). Exploring the chemistry and evolution of the isomerases. Proc. Natl. Acad. Sci. U S A *113*, 1796–1801.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, C.W. Carter, Jr., ed. (Academic Press), pp. 307–326.

Overkamp, K.E., Gasper, R., Kock, K., Herrmann, C., Hofmann, E., and Frankenberg-Dinkel, N. (2014). Insights into the biosynthesis and assembly of Cryptophycean phycobiliproteins. J. Biol. Chem. *289*, 26691–26707.

Palenik, B. (2001). Chromatic adaptation in MarineSynechococcus strains. Appl. Environ. Microbiol. *67*, 991–994.

Phillips, J.C., Hardy, D.J., Maia, J.D.C., Stone, J.E., Ribeiro, J.V., Bernardi, R.C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., et al. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. J. Chem. Phys. *153*, 044130.

Sanfilippo, J.E., Nguyen, A.A., Karty, J.A., Shukla, A., Schluchter, W.M., Garczarek, L., Partensky, F., and Kehoe, D.M. (2016). Self-regulating genomic island encoding tandem regulators confers chromatic acclimation to marine *Synechococcus*. Proc. Natl. Acad. Sci. U S A *113*, 6077–6082.

Sanfilippo, J.E., Garczarek, L., Partensky, F., and Kehoe, D.M. (2019a). Chromatic acclimation in cyanobacteria: a diverse and widespread process for optimizing photosynthesis. Annu. Rev. Microbiol. *73*, 407–433.

Sanfilippo, J.E., Nguyen, A.A., Garczarek, L., Karty, J.A., Pokhrel, S., Strnat, J.A., Partensky, F., Schluchter, W.M., and Kehoe, D.M. (2019b). Interplay between differentially expressed enzymes contributes to light color acclimation in marine *Synechococcus*. Proc. Natl. Acad. Sci. U S A *116*, 6457–6462.

Scheer, H., and Zhao, K.-H. (2008). Biliprotein maturation: the chromophore attachment: biliprotein chromophore attachment. Mol. Microbiol. *68*, 263–276.

Schirmer, T., Bode, W., and Huber, R. (1987). Refined three-dimensional structures of two cyanobacterial C-phycocyanins at 2.1 and 2.5 Å resolution. J. Mol. Biol. *196*, 677–695.

Schluchter, W.M., Shen, G., Alvey, R.M., Biswas, A., Saunée, N.A., Williams, S.R., Mille, C.A., and Bryant, D.A. (2010). Phycobiliprotein biosynthesis in cyanobacteria: structure and function of enzymes involved in post-translational modification. In Recent Advances in Phototrophic Prokaryotes, P.C. Hallenbeck, ed. (Springer New York), pp. 211–228.

Sheldrick, G.M. (2008). A short history of *SHELX*. Acta Crystallogr. A. *64*, 112–122.

Shen, G., Saunée, N.A., Williams, S.R., Gallo, E.F., Schluchter, W.M., and Bryant, D.A. (2006). Identification and characterization of a new Class of bilin lyase the cpcT gene encodes A bilin lyase responsible for attachment OF Phycocyanobilin to CYS-153 on the β-subunit of Phycocyanin in Synechococcus SP. PCC 7002. J. Biol. Chem. *281*, 17768–17778.

Shen, G., Schluchter, W.M., and Bryant, D.A. (2008). Biogenesis of Phycobiliproteins. I. cpcS-I and cpcU mutants of the cyanobacterium Synechococcus sp. PCC 7002 define a heterodimeric phycocyanobilin lyase specific for beta -phycocyanin and allophycocyanin subunits. J. Biol. Chem. *283*, 7503–7512.

Shukla, A., Biswas, A., Blot, N., Partensky, F., Karty, J.A., Hammad, L.A., Garczarek, L., Gutu, A., Schluchter, W.M., and Kehoe, D.M. (2012). Phycoerythrin-specific bilin lyase-isomerase controls blue-green chromatic acclimation in marine Synechococcus. Proc. Natl. Acad. Sci. *109*, 20136–20141.

Stumpe, H., Müller, N., and Grubmayr, K. (1993). The addition of methyl-2-mercaptoacetate to phycocyanobilin dimethyl ester: a model reaction for biliprotein biosynthesis? Tetrahedron Lett. *34*, 4165–4168.

Tu, J.-M., Zhou, M., Haessner, R., Plöscher, M., Eichacker, L., Scheer, H., and Zhao, K.-H. (2009). Toward a mechanism for biliprotein lyases: revisiting nucleophilic addition to phycocyanobilin. J. Am. Chem. Soc. *131*, 5399–5401.

Wagner, J.R., Brunzelle, J.S., Forest, K.T., and Vierstra, R.D. (2005). A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. Nature *438*, 325–331.

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

Yang, X., Stojković, E.A., Kuk, J., and Moffat, K. (2007). Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RpBphP3, reveals residues that modulate photoconversion. Proc. Natl. Acad. Sci. U S A *104*, 12571–12576.

Yang, X., Kuk, J., and Moffat, K. (2008). Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: photoconversion and signal transduction. Proc. Natl. Acad. Sci. U S A *105*, 14715–14720.

Zhao, C., Höppner, A., Xu, Q.-Z., Gärtner, W., Scheer, H., Zhou, M., and Zhao, K.-H. (2017). Structures and enzymatic mechanisms of phycobiliprotein lyases CpcE/F and PecE/F. Proc. Natl. Acad. Sci. U S A *114*, 13170–13175.

Zhao, K.-H., Deng, M.-G., Zheng, M., Zhou, M., Parbel, A., Storf, M., Meyer, M., Strohmann, B., and Scheer, H. (2000). Novel activity of a phycobiliprotein lyase: both the attachment of phycocyanobilin and the isomerization to phycoviolobilin are catalyzed by the proteins PecE and PecF encoded by the phycoerythrocyanin operon. FEBS Lett. *469*, 9–13.

Zhou, J., Gasparich, G.E., Stirewalt, V.L., de Lorimier, R., and Bryant, D.A. (1992). The cpcE and cpcF genes of Synechococcus sp. PCC 7002. Construction and phenotypic characterization of interposon mutants. J. Biol. Chem. *267*, 16138–16145.

Zhou, W., Ding, W.-L., Zeng, X.-L., Dong, L.-L., Zhao, B., Zhou, M., Scheer, H., Zhao, K.-H., and Yang, X. (2014). Structure and mechanism of the phycobiliprotein lyase CpcT. J. Biol. Chem. *289*, 26677–26689.

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli: DH5α	Invitrogen	CAT#:18-258-012
Escherichia coli: BL21(DE3)	Invitrogen	CAT#:C600003
Chemicals, peptides, and recombinant proteins		
Selenomethionine	Acros Organics	CAT#: 3211-76-5
TALON resin	Takara	CAT#: 635504
DL-Malic acid	Sigma Aldrich	LOT#: MKBW2323V
Synechococcus sp. A1562 MpeQ	This study	Genbank: AGW21721
Synechococcus sp. A1562 MpeW	This study	Genbank: AGW21717
Synechococcus sp. A1562 MpeA	This study	Genbank: AVH76704
See Table S2 for more recombinant proteins	This study	
Critical commercial assays		
GeneArt Seamless Cloning and Assembly Enzyme Mix	Invitrogen	CAT#: A14606
Deposited data		
MpeQ crystal structure in space group C2	This study	PDB:7MC4
MpeQ crystal structure in space group $C222_1$	This study	PDB:7MCH
Recombinant DNA		
pCDF Duet-1 vector	Novagen	CAT#:71340-3
pET Duet-1 vector	Novagen	CAT#:71146-3
Plasmid: A1562 Nus-MpeQ/pET44b	(Grébert et al., 2021)	N/A
Plasmid: A1562 Nus-MpeW/pET44b	(Grébert et al., 2021)	N/A
Plasmid: A1562 HT-MpeA/pCOLA	(Grébert et al., 2021)	N/A
Plasmid: NT-PebS/NT-HOI/pACYC	(Kronfel et al., 2019)	N/A
Plasmid: A1562 NT-MpeQ/pCDF	This study	N/A
Plasmid: A1562 HT-MpeQ/pCDF	This study	N/A
Plasmid: A1562 NT-MpeW/pET	This study	N/A
See Table S4 for more recombinant DNA		
Software and algorithms		
PISA	(Krissinel and Henrick, 2007)	https://www.ebi.ac.uk/pdbe/pisa
HKL2000	(Otwinowski & Minor, 1997)	https://hkl-xray.com/hkl-2000
XIA2 (CCP4)	(Winter, 2010)	https://www.ccp4.ac.uk/
Shelx97 (CCP4)	(Sheldrick, 2008)	https://www.ccp4.ac.uk/
Phaser (Phenix)	(McCoy et al., 2007)	http://www.phenix-online.org/
Phenix.refine	(Adams et al., 2010)	http://www.phenix-online.org/
Coot	(Emsley and Cowtan, 2004)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PyMOL	(DeLano, 2002)	https://pymol.org/2/
MUSCLE	(Edgar, 2004)	https://www.ebi.ac.uk/Tools/msa/muscle/
BioPython	(Cock et al., 2009)	https://biopython.org/
WebLOGO	(Crooks, 2004)	http://weblogo.threeplusone.com/
SwissMODEL	(Biasini et al., 2014)	https://swissmodel.expasy.org/
NAMD	(Phillips et al., 2020)	https://www.ks.uiuc.edu/Research/namd/

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

CellPress

RESOURCE AVAILABILITY

Lead contact

Xiaojing Yang (xiaojing@uic.edu).

Materials availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xiaojing Yang (xiaojing@uic.edu).

Data and code availability

- The coordinates and structure factor amplitudes of MpeQ in the space group C2 and $C222_1$ have been deposited in the Protein Data Bank under the accession numbers 7MC4 and 7MCH, respectively.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The *mpeQ* and *mpeW* genes from the CA-B strain *Synechococcus* A15-62 were cloned into the pCDF-Duet 1 and pET-Duet 1 vectors using the restriction sites SacI/NotI (or BglII/EcoRV) and NcoI/HindIII, respectively. MpeQ carrying a N-terminal 6xHis affinity tag was over-expressed in *E. coli* BL21(DE3) by isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) induction followed by overnight shaking at 18°C.

METHOD DETAILS

Protein purification and site-directed mutagenesis

After cell harvesting and cell lysis, His-tagged MpeQ protein was extracted by Co^{2+}-affinity chromatography and further purified by anion exchange chromatography (HiTrap Q- HP column). The selenomethionine (SeMet)-derivatized MpeQ protein was prepared according to the standard protocols (Doublié, 1997) and purified using the same protocols as for the native MpeQ.

Site-directed mutagenesis was carried out using combined overlapping polymerase Chain Reaction (COE-PCR) as previously described with modifications (Hussain and Chong, 2016). Mutagenic primers were created using the Thermo Fisher Scientific GeneArt primer and construct design tool (https://www.thermofisher.com/order/oligoDesigner/mutagenesis) or NEB primer design tool (http://nebasechanger.neb.com) (Tables S2 and S3). Quintuple mutants were created using the GeneArt Seamless Cloning and Assembly Enzyme Mix (Invitrogen). All mutants of MpeQ were over-expressed and purified using the same protocols as for the wild type (Table S4).

Crystallization, data collection and structure determination

Purified MpeQ protein was crystallized using the hanging drop vapor diffusion method by mixing the protein sample (5 mg/mL) with the crystallization solution (2.1 M DL-Malic acid, pH 7.0) in 1:1 ratio. Single crystals obtained by macro-seeding were cryoprotected in the mother liquor containing 25% glycerol for X-ray diffraction experiments. All diffraction images were collected at the Life-Science Consortium Access Team (LS-CAT) beam stations at the Advanced Photon Source, Argonne National Laboratory. The diffraction datasets were indexed, integrated and scaled using HKL2000 (Otwinowski and Minor, 1997) and/or xia2 in CCP4 (Winter, 2010).

The crystal structure of SeMet-MpeQ was determined using the SAD method. The initial SAD phasing was carried out using the MpeQ dataset of 2.9 Å resolution in the space group of $C222_1$ (Shelx97) (Sheldrick, 2008). An initial model was used as a search model to determine the MpeQ structure in the space group C2 by molecular replacement (Phaser) (McCoy et al., 2007). The final model was refined at 2.5 Å resolution with the *R*-factor and free *R*-factor of 0.232 and 0.285, respectively (Phenix.refine) (Adams et al., 2010). The final model in the space group $C222_1$ was refined at 2.95 Å resolution with the *R*-factor and free *R*-factor of 0.245 and 0.307, respectively. Model building and molecular graphics were done using Coot (Emsley and Cowtan, 2004) and/or PyMOL(DeLano, 2002).

Sequence analysis and identification of switch residues

Sequence alignment of 106 bilin lyases and lyase-isomerases in the MpeQWYZ family were aligned using MUSCLE (Edgar, 2004) (Dataset S1). Each sequence was assigned a lyase or lyase-isomerase function based on the published biochemical data (Grébert et al., 2021; Sanfilippo et al., 2019b; Shukla et al., 2012). Positions identical or differing between lyases and isomerases were identified using Biopython 1.77 (Cock et al., 2009) (Dataset S2). Sequence logo was generated using WebLogo (Crooks, 2004).

Enzyme activity assays based on a co-expression system

A three-plasmid co-expression system was used to test the lyase activities of MpeQ and MpeW. Specifically, MpeQ or MpeW was co-expressed in *E. coli* BL21(DE3) cells with MpeA (pCOLA-Duet vector) along with a pACYC-Duet vector carrying heme oxygenase

Please cite this article in press as: Kumarapperuma et al., Crystal structure and molecular mechanism of an E/F type bilin lyase-isomerase, Structure (2022), https://doi.org/10.1016/j.str.2022.01.007

(HO1) and PEB synthase (PebS) genes (Kronfel et al., 2019). His-tagged MpeQ or non-tagged MpeW was then co-purified with MpeA using Co2+-affinity column (Talen). Eluted fractions were then examined by SDS polyacrylamide electrophoresis (SDS-PAGE), absorption spectroscopy, and/or fluorescence spectroscopy. The bilin lyase activity was measured by the chromophorylated MpeA detected either by Zn-fluorescence (Berkelman and Lagarias, 1986), absorption spectroscopy (Shimadzu UV-2600 UV-Vis spectrophotometer), and/or fluorescence spectroscopy (Perkin Elmer LS55 fluorescence spectrophotometer). MpeA-PUB and MpeA-PEB have characteristic absorption peaks at 495 nm and 550 nm, respectively, while their corresponding emission peaks are around 500 nm and 565 nm. The lyase activities of the wild type MpeQ and mutants were quantified and compared by the absorbance readings at 495 nm normalized by the corresponding MpeA band intensities detected by SDS-PAGE (ImageJ). The relative activity of a specific mutant (in percentage) was calculated in comparison to the wild type for which the lyase activity was set to 100%.

Molecular dynamics (MD) simulations

72-ns MD simulations were performed on the structures of apo-MpeA and MpeQ (PDB ID: 7MC4). The starting structure for apo-MpeA was based on a homology model of MpeA obtained by SwissModel (Biasini et al., 2014) with all the pigments removed. All minimization and MD simulation steps were performed using NAMD (Phillips et al., 2020) with CHARMM36m force field (Huang et al., 2017). The simulation was performed on 32 processors of a Cray T3E parallel supercomputer at University of Illinois Chicago High-Performance Computing. Frames were collected at 100-ps intervals for the simulation length of 72 ns, giving 720 frames for conformational analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical data in Table 1 for the crystallography data and structural analysis were obtained from the outputs of HKL2000 (Otwinowski and Minor, 1997) and Phenix (Adams et al., 2010).