Title

Evolutionary analysis of the functional plasticity of *Staphylococcus aureus* C₃₀ carotenoid synthase

Maiko Furubayashi¹, Kyoichi Saito¹ and Daisuke Umeno^{1,2*}

¹Department of Applied Chemistry and Biotechnology, Chiba University, 1-33 Yayoi-cho, Inage, Chiba, 263-8522, Japan

²Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan.

Running Title

Directed evolution of carotenoid synthase

Keywords

carotenoid, size specificity, directed evolution, carotene synthase, screening

* Corresponding author. E-mail: umeno@faculty.chiba-u.jp, Tel/Fax: +81-43-290-3413

1 **Abstract** (<250 words)

2
3 Most natural carotenoids have 40-carbon (C₄₀) backbones, while some bacteria produce
4 carotenoids with C₃₀ backbones. Carotenoid backbone synthases, the enzyme that catalyze the
5 first committed step in carotenoid biosynthesis, are known to be highly specific. Previously,
6 using C₃₀ backbone synthase (diapophytoene synthase, CrtM) from *Staphylococcus aureus*, we
7 reported two size-shifting mutations, F26A and W38A, which confer C₄₀ synthase activity at the
8 cost of the original C₃₀ synthase activity. In this study, we performed a directed evolution of the
9 “C₄₀ specialist” variant CrtM_{F26A} in search of mutations that restore the original C₃₀ synthase
10 function. Examination of the resultant mutants, together with the site-directed mutagenesis
11 study identified three new mutations (H12A, D27A and I240F) that affect the size specificity of
12 this enzyme. After re-defining the reading frame, we obtained CrtM variants that are highly
13 active in C₃₀ and C₄₀ carotenoid synthesis.
14

1 **Introduction**

2 Carotenoids are widespread natural pigments found in various organisms, and they play
3 important biological roles as light-harvesting components (1), photoprotectants (2), antioxidants
4 (3), and membrane fluidity controllers (4,5). Most natural carotenoids have 40-carbon (C_{40})
5 backbones, while some bacteria produces carotenoids with C_{30} backbones. To date, C_{30}
6 carotenoids and C_{40} carotenoids have not been found to be produced in the same organism; C_{30}
7 carotenoids have never been found in C_{40} carotenoid producers, and C_{40} carotenoids have never
8 been found in C_{30} carotenoid producers (6). This is surprising, as most of the carotenoid
9 modification enzymes tested are promiscuous, and they accept various (C_{30} , C_{35} , C_{40} , C_{45} , C_{50})
10 backbone sizes (6).

11 The first committed step in the carotenoid pathway is the condensation of two
12 isoprenyl diphosphates to synthesize a carotenoid backbone. In C_{40} carotenoid synthesis, two
13 geranylgeranyl diphosphates (C_{20} PPs) are selectively condensed by phytoene synthase to yield
14 phytoene (C_{40} backbone) (Fig. 1). In C_{30} carotenoid synthesis, two molecules of farnesyl
15 diphosphates (C_{15} PPs) are converted to diapophytoene (C_{30} backbone). This reaction step is
16 highly specific, and it is responsible for the overall size specificity of the carotenoid pathway.
17 We are interested in how carotenoid backbone synthases achieve size specificity for their
18 respective reactions (6,7). Previously, we reported two size-shifting mutations, F26A and
19 W38A, that confer C_{40} synthase activity to C_{30} backbone synthase (diapophytoene synthase,
20 CrtM) in *Staphylococcus aureus* (8). Both mutations were reported to not only confer C_{40}
21 synthase activity but also decrease the native C_{30} synthase function, resulting in specialist
22 variants with C_{40} synthase function. Many laboratory-evolved enzymatic functions are
23 promiscuous (9,10), yielding ‘generalist’ variants that possess both their original and novel
24 activities. In this study, we sought to determine whether it is possible for CrtM to become
25 C_{30}/C_{40} generalist enzymes.

1 Based on the structure of the homolog protein human squalene synthase (11) and site
2 saturation mutagenesis, we reported that the F26 and W38 residues reside within the reaction
3 cavity of CrtM, and substitution of these residues with smaller amino acids enables CrtM to
4 accommodate larger substrates by enlarging the reaction cavity of this enzyme (8). The crystal
5 structure of the F26A mutant (12) confirmed our original speculation that this mutation removes
6 the steric block against larger substrates, enabling the condensation of larger substrates.
7 Furthermore, this structure also revealed that the W38 residue is located on the surface of the
8 protein (see below) and was not directly involved in substrate recognition.

9 In this study, we performed directed evolution of the “C₄₀ specialist” variant,
10 CrtM_{F26A}, in search of mutations that improve (restore) the original C₃₀ function. We identified a
11 mutation that will produce a “C₃₀ specialist” CrtM variant while retaining the F26A mutation.
12 *Additionally, directed evolution restored the reading frame of crtM gene; with acquiring*
13 *additional six-amino acids in the C-terminus, the F26A and W38A mutant of CrtM turned to*
14 *become C₃₀/C₄₀ generalist.* Site directed mutagenesis indicated that the W38A mutation
15 probably loosens the rigidity of the enzyme, allowing for the condensation of substrates with
16 various sizes.

Materials and Methods

Bacterial strains and genetic manipulations. *E. coli* XL10-Gold Kan^R (Tet^r $\Delta(mcrA)183$ $\Delta(mcrCB-hsdSMR-mrr)173$ *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lacI^qZAM15 Tn10* (Tet^r) Tn5 (Kan^r) Amy]) (Stratagene, CA) was used for DNA cloning and carotenoid colony screening, and XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^qZAM15 Tn10* (Tet^r)]) (Stratagene, CA) was used for carotenoid production analysis in liquid media. Polymerase chain reaction (PCR) amplifications were performed using *Vent* polymerase for cloning and *Taq* polymerase for mutagenic PCR (both from New England BioLabs, MA). Restriction enzymes and DNA modifying enzymes were purchased from New England BioLabs (MA). DNAs were purified using the Zymo DNA Clean and Concentrator (Zymo Research, CA). Chemically competent cells were prepared using the Z-Competent™ *E. Coli* Transformation Kit and Buffer Set (Zymo Research, CA). Plasmid DNA was prepared using the Plasmid Miniprep Kit from Sigma-Aldrich. Nucleotide sequences were confirmed with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, CA) and analyzed using an ABI 3130 Genetic Analyzer (Applied Biosystems, CA).

Plasmids. pUC-*crtM* and pUC-*crtB* were previously constructed (7) by directionally inserting *S. aureus crtM* or *Pantoea ananatis crtB* into the *XbaI/XhoI* restriction sites in the pUC18m vector (7). pUC-*crtM* variants were made by whole plasmid inverse PCR, followed by a FASTR (13) reaction (type IIS restriction enzyme cloning). The codons adopted for each mutagenesis reaction were as follows: GCG for F26A, GCG for W38A, TTT for I240F, GCG for H12A, and GCG for D12A. For the construction of pUC-*fl-crtM* (full-length *crtM*), adenine at the 839th position was deleted from the original pUC-*crtM* construct using the following primers to PCR-amplify the *crtM* gene: Fwd, 5'-CTTTACACTTTATGCTTCCGG-3'; Rev, 5'-TTTTCTCGAGCTATATTCTATGATATTTACTATTTATTTTCATGAAAC-3'. The

resultant PCR fragment was cloned into the *XbaI/XhoI* restriction sites in the pUC18m vector (7). pAC-*crtE-crtI* (8) and pAC-*crtN* (8), which are based on the pACmod vector (pACYC184-derivative), encode CrtE, CrtI or CrtN, respectively, constitutively expressed by the leaky expression of the *lac* promoter.

Construction of the CrtM_{F26A} library. Using pUC-*crtM_{F26A}* as a template, error prone PCR was performed by the addition of 10 μ M or 50 μ M Mn²⁺ (14). Amplified PCR fragments were cloned into the *XbaI/XhoI* restriction sites in the pUC18m vector (7), and the resulting constructs were transformed into XL10-Gold Kan^R chemically competent cells. A portion of the transformants were plated on LB-Lennox agar plate to assess transformation efficiency, and the rest were cultured in 10 mL LB-Lennox containing 50 μ g/mL carbenicillin (carb) at 37°C for 12 h. Two milliliters of liquid culture was collected to purify the plasmid mixture. The library sizes of the resultant plasmid pool were 7.7×10^5 (10 μ M Mn²⁺) and 2.0×10^5 (50 μ M Mn²⁺).

C₃₀ synthase activity screening of CrtM variants. The plasmid libraries described above were transformed into XL10-Gold Kan^R harboring pAC-*crtN-idi* and plated on an LB-Lennox agar plate containing 50 μ g/mL carb and 30 μ g/mL chloramphenicol (cm). After incubation at 37°C for 12 h, the formed colonies were lifted using nitrocellulose membranes (BioTrace™ NT Nitrocellulose Transfer Membrane, Pall Corporation, NY), transferred onto a fresh LB-Lennox (carb/cm) agar plate, and incubated at room temperature (approximately 25°C) for an additional 24 h. Each colony exhibited a different level of pigmentation. Colonies with a deeper yellow color (compared to colonies expressing CrtM_{F26A}) were isolated, and approximately 1500 colonies in total were screened and streaked onto a new LB-Lennox (carb/cm) agar plate topped with a nitrocellulose membrane. The 10 clones isolated from the streak that displayed a deep yellow hue were minipreped and re-transformed into fresh XL1B-Blue harboring pAC-*crtN* to

confirm their colony hue. In all, 4 clones were isolated.

Analysis of C₃₀- and C₄₀- synthase activity. pUC-*crtM* and its variants were transformed into XL1-Blue harboring pAC-*crtN* to score C₃₀ synthase activity (8) or XL1-Blue harboring pAC-*crtE-crtI* to score the C₄₀ synthase activity (8). In both cases, transformants were plated onto LB-Lennox (carb/cm) agar plates to form colonies. These colonies were picked and inoculated into 500 µL LB-Lennox (carb/cm) media in a 96-well deep well plate and cultured at 37°C, 1000 rpm for 16 h. An aliquot (40 µL) of these pre-cultures was transferred to 2 mL Terrific Broth (TB) (carb/cm) in a 48-well deep well plate and cultured at 30°C, 1000 rpm for 48 h. Cells were harvested, washed with saline, and centrifuged to obtain cell pellets; the supernatants were discarded. After briefly vortexing, 1 mL acetone was added to each of the cell pellets, and they were immediately vortexed for 1 min to extract carotenoids, followed by centrifugation. The absorbance spectra (350-650 nm at 5 nm interval) were analyzed for acetone extracts using SpectraMax Plus³⁸⁴ (Molecular Devices, CA).

The pigmentation level of each culture was determined from the lambda max (470 nm for C₃₀ carotenoids and 475 nm for C₄₀ carotenoids) of the resultant extract, using the molar adsorption coefficients of diaponeurosporene (147,000 M⁻¹cm⁻¹) or lycopene (185,000 M⁻¹cm⁻¹).

Results

Directed evolution of CrtM_{F26A} for elevated C₃₀ synthase function.

Previously, we reported that a single amino acid substitution, F26A, confers C₄₀ carotenoid (phytoene) synthase function to CrtM, while it greatly decreases its original function as a C₃₀ carotenoid (diapophytoene) synthase (8). As a consequence, CrtM_{F26A} performed in the same manner as a specialized C₄₀ carotenoid synthase (“C₄₀ specialist”) in *E. coli*. To search for mutations that restore the C₃₀ activity of CrtM_{F26A}, we performed directed evolution of CrtM_{F26A} to obtain mutations that confer higher C₃₀ synthase activity. Note that the codon assigned for Ala 26 in the *crtM_{F26A}* gene is GCG, which was chosen so that a single nucleotide change could not restore the Phe codon (TTN, N=A/G/T/C). Random mutations were introduced into the entire *crtM_{F26A}* gene using error-prone PCR. The resultant CrtM_{F26A} library was co-transformed with pAC-*crtN*. Here, the production of C₃₀ carotenoid (diapophytoene) is visualized using the diapophytoene desaturase (CrtN) from *S. aureus* (Fig. 2A). We originally isolated 28 clones with a deeper yellow hue from approximately 1500 colonies. Streaking on fresh agar plates resulted in the selection of 10 clones that were clearly deeper in color. Plasmids from each of these 10 clones were minipreped and re-transformed into XL1-Blue cells harboring pAC-*crtN*. Finally, four variants (named M1-M4) with the highest color intensity were identified. All these CrtM mutants also showed improved C₃₀ synthase activity in liquid media (Fig. 2B).

To evaluate the C₃₀ and C₄₀ synthase activities of the pUC-*crtM* variants, we used the two-plasmid system that we developed previously (8) (Fig. 3). Briefly, C₃₀ carotenoid synthase activity was selectively scored by coexpression of pUC-*crtM* variants with pAC-*crtN*, where CrtN selectively converts C₃₀ carotenoid (diapophytoene) into diaponeurosporene, a yellow-colored carotenoid. The C₄₀ synthase activity of the CrtM variants was selectively scored by coexpression of pUC-*crtM* variants with pAC-*crtE-crtI*. Here, CrtE (geranylgeranyl diphosphate synthase) supplies geranylgeranyl diphosphate (C₂₀PP), the precursor for C₄₀

1 carotenoid. CrtI (phytoene desaturase) converts the C₄₀ carotenoid backbone (phytoene) into
2 lycopene, a red-colored pigment. Although CrtI has weak desaturase activity toward C₃₀
3 carotenoids (7,15), the co-expression of CrtE significantly reduces the cellular availability of
4 C₁₅PP for C₃₀ carotenoid synthesis, resulting in almost complete abolishment of C₃₀ synthesis,
5 even for the wildtype CrtM (7). In combination with the substrate preference of CrtI for its
6 cognate (C₄₀) substrates, the pigmentation of cells harboring pAC-*crtE-crtI* is very selective for
7 C₄₀ carotenoid (phytoene) formation.

8 By measuring the pigmentation in both contexts, we scored the C₃₀ and C₄₀ synthase
9 activities of the 4 variants. When plotted in a 2-dimensional graph, where the *x*-axis represents
10 the new function (C₄₀ synthase function) and the *y*-axis represents the original function (C₃₀
11 synthase function) (Fig. 4), wildtype CrtM is located on the left-top corner (“C₃₀-specialist”
12 area), while CrtB (native C₄₀ synthase from *P. ananatis*) and CrtM_{F26A} are located in the
13 right-bottom corner (“C₄₀-specialist” area). CrtM_{F26A} displayed minimal C₃₀ (original) function
14 and behaved more like CrtB. M1, M3 and M4 behaved more like “C_{30/40}-generalists,” displaying
15 a subtle increase in C₃₀ synthase activity without losing C₄₀ activity. In contrast, M2 appeared to
16 be a “C₃₀-specialist”; it showed the largest increase in C₃₀ activity, reaching a level that was
17 comparable to wildtype CrtM, while the C₄₀ activity was significantly decreased.

18 19 **Mutational analysis of CrtM variants with elevated C₃₀ function.**

20 The “C₃₀-specialist” variant M2 had two amino acid substitutions, V133A and I240F, in
21 addition to the original mutation, F26A (**Table 1**). Based on the homology model of the enzyme
22 human squalene synthase (8), we have previously proposed that the F26A mutation enlarged the
23 reaction pocket of CrtM, allowing it to accommodate C₂₀PP. Subsequently, structural
24 determination of CrtM and its variants by Lin *et al.* confirmed our hypothesis (12). Based on the
25 mapped structure on this CrtM structure (12), I240F, which is located on the helix neighboring

F26A, was likely responsible for the phenotypic change (restoration of C₃₀ function) of CrtM_{F26A} (Fig. 5). We assumed that the substitution of I240 with larger amino acids (such as Phe) would decrease of the substrate cavity size, making the enzyme prefer reactions with smaller substrates. Indeed, the addition of a single mutation, I240F, to CrtM_{F26A} was sufficient to confer a similar phenotype (C_{30/40} function) with M2; CrtM_{F26A/I240F} performed as a more specific C₃₀ synthase than M2 (Fig. 5). Although I240F ablated the C₄₀ synthase activity conferred by F26A, this mutation alone rather slightly increased the C₄₀ synthase activity (CrtM_{I240F}).

Restoration of the reading frame improved total CrtM activity.

Three mutants (M1, M3 and M4) shared a single nucleotide deletion near the C-terminus of the gene (**Table 1**). Our parent pUC-*crtM* plasmid contained the natural *S. aureus* SA113 (ATCC 35556) genome sequence after the stop codon of *crtM*; therefore, these frame shift mutations detoured the original stop codon and extended the reading frame by 6 amino acids compared to the parent *crtM* gene (Fig. S1). Given this result, we searched for the available *S. aureus crtM* gene sequence in the database, realizing that the frame with six additional residues (total 287 aa) was the correct reading frame for CrtM. The first reported *crtM* gene sequence (16) from *S. aureus* Newman (EMBL accession no. X73889) had insertions at nucleotide 756 and 766, which putatively codes for a 255-aa (amino acid) protein (Fig. S1). The CrtM we have used throughout was the product of PCR-amplification from the genomic DNA from *S. aureus* SA03 (ATCC 35556), and it had an insertion at 839 compared to the current database, resulting in a 281 aa protein (Fig. S1). The *S. aureus* genome used as the template for PCR in previous reports, including ours, might possess mutations that produce *crtM* genes of different lengths. There are many out-frame stop codons near the C-terminus of the *crtM* gene sequence, but alignment of recent genome sequencing data from several available *S. aureus* genomic sequences clearly

1 indicate that the correct reading frame for this *crtM* gene has 287 residues. It should be noted that
2 both the parental and the evolved *crtM* gene were defined using the amber (TAG) stop codons.
3 Considering that the host strain XL1-Blue possesses *supE44*, an amber suppressor mutation, there could be
4 non-negligible occurrence of read-through at these positions until they meet the next stop codon. For the
5 parent (truncated) *crtM*, ochre (TAA) stop codon locates right after the amber stop codon (Fig. S1). For
6 the evolved (full-length) *crtM*, opal (TGA) stop codon appears 18 nucleotides (6 amino acids) after the
7 amber stop codon (Fig. S1).

8 We synthesized CrtM variants with this new reading frame and named them
9 FL-CrtMs (FL for Full Length). We compared the C₃₀ and C₄₀ synthase activities of wildtype
10 CrtM, CrtM_{F26A}, and CrtM_{W38A} with both the original (281 aa) and new (full length, 287 aa)
11 reading frames. The C₃₀ and C₄₀ synthase activities of FL-CrtM_{F26A} and FL-CrtM_{W38A} were
12 higher than those of the original (truncated) CrtM_{F26A} and CrtM_{W38A} (Fig. 6). It is possible that
13 our original CrtMs might be less stable than the new (full) frame variants; by increasing the
14 stability with an additional six amino acids in the C-terminus, both the F26A and W38A
15 mutants showed higher C₃₀ and C₄₀ synthase activities in FL-CrtM, approaching the level of
16 individual natural specialist synthases, FL-CrtM/CrtB. In sharp contrast, the C_{30/40} synthase
17 activity of the wildtype (mutation-free) CrtM was almost identical to those of FL-CrtM,
18 revealing the remarkable specificity of wildtype-CrtM against the non-cognate function as a C₄₀
19 synthase.

21 **W38A makes CrtM promiscuous due to the loss of inter-helix interactions.**

22 In our previous study (8), we reported that the W38A mutation makes CrtM a C₄₀-specific
23 synthase by “removing the bulky residue” in the substrate cavity, enabling the accommodation
24 of larger substrates. However, the C_{30/40} generalist-like behavior of FL-CrtM_{W38A} (Fig. 6) made
25 us re-map this mutation (W38A) on the recently solved structure of CrtM (12). Interestingly, the

W38 residue was not located inside the reaction cavity of CrtM, contrary to our original assumption based on the structure of the related enzyme, squalene synthase (11). Instead, the W38 residue is located on the surface of the CrtM protein (Fig. 7).

Next, we investigated how this mutation enables CrtM to perform as both C₃₀ and C₄₀ synthases. Closer examination of the crystal structure of CrtM indicated that the D27 and H12 residues are located very near W38 (Fig. 7), possibly interacting each other. To examine the effect of eliminating the possible interaction among these residues, we mutated D27 and H12 to Ala. Indeed, we observed a non-negligible increase in C₄₀ activity of the CrtM_{H12A}, CrtM_{D27A}, and CrtM_{H12A,D27A} mutants, without compensation in C₃₀ synthase activity (Fig. 7). We speculate that these interactions are important in “bundling” the two helices (helix A and B) that make up the sidewall of the substrate cavity. When mutations W38A, H12A, or D27A were introduced to disrupt these interactions, the substrate cavity acquired the flexibility needed to accommodate substrates with different sizes.

Note that the H12R or H12Q mutants were enriched in our first round of directed evolution of CrtM for C₄₀ synthase function (7,17). We did not notice the possible effect of the H12 mutation because of the prominent effects of mutations at W38A or E180G. Given the result in Fig. 7 and the observation that the W38C/H12Q mutant has slightly higher C₄₀ synthase activity than W38C (17), we propose that removing the interaction among H12-D27-W38 increases the flexibility of the reaction cavity of CrtM, thereby creating multi-talented carotenoid synthases.

Discussion

C₄₀ and C₃₀ carotenoid pathways are very similar except in the sizes of their backbones, and it is clear that they diverged from a common ancestral pathway. Nevertheless, no C₃₀ carotenoids are known to be C₄₀ carotenoid producers, and no C₄₀ carotenoids are detected in C₃₀ carotenoid

1 producers (6). We are interested in how this high level of “size-specificity” of carotenoid
2 synthases is achieved (6-8,17,18). We reported previously that CrtM (C₃₀ synthase) could easily
3 obtain C₄₀ synthase function with a single mutation, either F26A or W38A. For these two
4 mutations, a negative trade-off for C₃₀ and C₄₀ synthase activities was observed, and no other
5 C_{30/40} generalist variant was obtained. However, in general, it has been proposed that new
6 functions of an enzyme are usually acquired as a promiscuous function in addition to the
7 original function. The apparent trade-off between C₃₀/C₄₀ functions in CrtM variants has
8 hampered the creation of ‘generalist’ CrtM variants with C_{30/40}-functions.

9 In this work, by evolving the F26A mutant of CrtM, we created ‘generalist’ mutants
10 that possess C₃₀/C₄₀ functions. We isolated a single mutation (I240F) that cancels the
11 specificity-shifting effect of F26A. Interestingly, this mutation alone slightly increases C₄₀
12 synthase activity, indicating that the remarkable specificity of wildtype CrtM for C₃₀ synthase
13 function (or strong specificity *against* C₄₀ function) is due to the delicate shape management of
14 multiple residues residing in the reaction cavity. Additionally, by correcting the reading frame,
15 both F26A and W38A mutations, which were originally reported as specificity shifting
16 mutations (8), produced ‘generalist’ carotenoid synthase. Mutation analysis of surrounding
17 residues (H12 and D27) indicated that W38 contributes to the size specificity of CrtM by
18 making the reaction pocket more rigid. Thus, CrtM can possess both C₃₀ and C₄₀ carotenoid
19 activities with at least two completely different mechanisms.

20 Natural C₄₀ synthase CrtB from *P. ananatis* does not display any C₃₀ synthase
21 activity (Fig. 6), as demonstrated in a previous report (7,15). In contrast, some level of C₃₀
22 synthase activity always accompanies the C₄₀ synthase activity of CrtM variants. In general, to
23 make a totally specific variant with novel function, many additional amino acid substitutions are
24 needed (19,20). It remains unknown how many evolutionary steps (amino acid substitutions) are
25 required to maintain C₄₀ synthase function and lose C₃₀ synthase function. Similarly, it is

- 1 unclear whether C_{40} synthase CrtB (natural C_{40} synthase) can acquire C_{30} synthase function and
- 2 become a specific C_{30} synthase. Further rounds of laboratory evolution of carotenoid synthases
- 3 might answer these questions.

1 **Acknowledgements**

2 D.U. is supported by the PRESTO (JST) and the MEXT/JSPS Grant-in-Aid for Scientific
3 Research on Innovative Areas 23108507. M.F. is supported by a JSPS fellowship for young
4 scientists. This work is financially supported by the Mishima-Kaiun Memorial Foundation, the
5 Futaba Electronics Memorial Foundation, and the Shorai Foundation for Science and
6 Technology.

References

1. Polivka, T., Frank, H. A.: Molecular factors controlling photosynthetic light harvesting by carotenoids. *Acc. Chem. Res.*, **43**, 1125-1134 (2010).
2. Demmig-Adams, B., Adams, W. W.: Photoprotection and Other Responses of Plants to High Light Stress. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 599-626 (1992).
3. Edge, R., McGarvey, D. J., Truscott, T. G.: The carotenoids as anti-oxidants—a review. *J. Photochem. Photobiol. B.*, **41**, 189-200 (1997).
4. Mishra, N. N., Liu, G. Y., Yeaman, M. R., Nast, C. C., Proctor, R. A., McKinnell, J., Bayer, A. S.: Carotenoid-related alteration of cell membrane fluidity impacts *Staphylococcus aureus* susceptibility to host defense peptides. *Antimicrob. Agents Chemother.*, **55**, 526-531 (2011).
5. Wisniewska, A., Widomska, J., Subczynski, W. K.: Carotenoid-membrane interactions in liposomes: effect of dipolar, monopolar, and nonpolar carotenoids. *Acta Biochim. Pol.*, **53**, 475-484 (2006).
6. Umeno, D., Tobias, A. V., Arnold, F. H.: Diversifying carotenoid biosynthetic pathways by directed evolution. *Microbiol. Mol. Biol. Rev.*, **69**, 51-78 (2005).
7. Umeno, D., Tobias, A. V., Arnold, F. H.: Evolution of the C₃₀ carotenoid synthase CrtM for function in a C₄₀ pathway. *J. Bacteriol.*, **184**, 6690-6699 (2002).
8. Umeno, D., Arnold, F. H.: Evolution of a pathway to novel long-chain carotenoids. *J. Bacteriol.*, **186**, 1531-1536 (2004).
9. Matsumura, I., Ellington, A. D.: In vitro evolution of β -glucuronidase into a β -galactosidase proceeds through non-specific intermediates. *J. Mol. Biol.*, **305**, 331-339 (2001).
10. Aharoni, A., Gaidukov, L., Khersonsky, O., Mc, Q. G. S., Roodveldt, C., Tawfik, D. S.: The 'evolvability' of promiscuous protein functions. *Nat. Genet.*, **37**, 73-76 (2005).
11. Pandit, J., Danley, D. E., Schulte, G. K., Mazzalupo, S., Pauly, T. A., Hayward, C. M., Hamanaka, E. S., Thompson, J. F., Harwood, H. J., Jr.: Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis. *J. Biol. Chem.*, **275**, 30610-30617 (2000).
12. Lin, F. Y., Liu, C. I., Liu, Y. L., Zhang, Y., Wang, K., Jeng, W. Y., Ko, T. P., Cao, R., Wang, A. H., Oldfield, E.: Mechanism of action and inhibition of dehydrosqualene synthase. *Proc. Natl. Acad. Sci. USA*, **107**, 21337-21342 (2010).
13. Kotera, I., Nagai, T.: A high-throughput and single-tube recombination of crude PCR products using a DNA polymerase inhibitor and type IIS restriction enzyme. *J.*

- Biotechnol., **137**, 1-7 (2008).
14. **Cadwell, R. C., Joyce, G. F.:** Mutagenic PCR. PCR Methods Appl., **3**, S136-140 (1994).
15. **Raisig, A., Sandmann, G.:** Functional properties of diapophytoene and related desaturases of C₃₀ and C₄₀ carotenoid biosynthetic pathways. Biochim. Biophys. Acta, **1533**, 164-170 (2001).
16. **Wieland, B., Feil, C., Gloria-Maercker, E., Thumm, G., Lechner, M., Bravo, J. M., Poralla, K., Gotz, F.:** Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol., **176**, 7719-7726 (1994).
17. **Umeno, D., Hiraga, K., Arnold, F. H.:** Method to protect a targeted amino acid residue during random mutagenesis. Nucleic Acids Res., **31**, e91 (2003).
18. **Umeno, D., Arnold, F. H.:** A C₃₅ carotenoid biosynthetic pathway. Appl. Environ. Microbiol., **69**, 3573-3579 (2003).
19. **Fasan, R., Meharena, Y. T., Snow, C. D., Poulos, T. L., Arnold, F. H.:** Evolutionary history of a specialized p450 propane monooxygenase. J. Mol. Biol., **383**, 1069-1080 (2008).
20. **Tokuriki, N., Jackson, C. J., Afriat-Jurnou, L., Wyganowski, K. T., Tang, R., Tawfik, D. S.:** Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. Nat. Commun., **3**, 1257 (2012).

Figure and Table legends

FIG. 1. Carotenoid biosynthetic pathways. C₃₀ and C₄₀ carotenoid backbone pathways exist in nature. These pathways can be reconstructed in *E. coli* from endogenous C₁₅PP. In this study, we used C₃₀ carotenoid pathway enzymes (CrtM and CrtN) from *S. aureus* and C₄₀ carotenoid pathway enzymes (CrtE, CrtB, and CrtI) from *P. ananatis*.

FIG. 2. Directed evolution of CrtM_{F26A}. (A) Schematic procedure of the directed evolution experiment. Details are described in the Materials and Methods section. (B) C₃₀ carotenoid production of CrtM_{WT}, CrtM_{F26A}, and its variants. Cells co-expressing CrtMs with CrtN were cultured for 48 h in TB medium at 30°C. Carotenoids were acetone-extracted to determine pigment production in a diaponeurosporene equivalent. Bars heights represent the average of four replicates, and error bars represent the standard deviation of four experiments.

FIG. 3. Scoring systems for C₃₀ and C₄₀ synthase activities. (A) The system used to measure C₃₀ synthase function. CrtN is expressed from the *lac* promoter on the pACmod vector. C₃₀ backbone (diapophytoene) synthesis by CrtM variants (on a pUC18m vector) can be visualized by conversion to diaponeurosporene, a yellow pigment. (B) The system used to measure C₄₀ synthase function. CrtE and CrtI are expressed under the *lac* promoter on the pACmod vector. CrtE synthesizes C₂₀PP from endogenous C₁₅PP to supply substrate for C₄₀ carotenoid synthesis. The synthesis of C₄₀ backbone (phytoene) is visualized by its conversion into lycopene, a red pigment.

FIG. 4. C₃₀ and C₄₀ synthase activity of isolated CrtM variants. C₃₀ and C₄₀ activities were determined as described in Fig. 3. Data points represent the averages of four replicates, and error bars represent the standard deviation.

FIG. 5. “C₃₀-specific” variant M2. (A) Residues F26 and I240 are mapped on the structure of CrtM (PDB ID: 2ZCP). The helix consisting of amino acids 1-25 and 266-287 was hidden to provide a better view of the relevant residues. (B) C₃₀ and C₄₀ synthase activities of CrtM I240F variants. C₃₀ or C₄₀ synthase functions of pUC-*crtM* variants or *crtB* were determined as indicated in Fig. 3. Data points represent the average of four replicates, and error bars represent the standard deviation.

FIG. 6. C₃₀ and C₄₀ synthase activities of CrtM variants with full (287 aa) or truncated (281 aa) reading frames. The C₃₀ or C₄₀ synthase functions of pUC-*crtM* variants (or *crtB*) were determined as indicated in Fig. 3. Data points represent the average of four replicates, and error bars represent the standard deviation.

FIG. 7. Effect of mutations surrounding residue 38. (A) Residues H12, D27 and W38 mapped onto the crystal structure of CrtM (PDB ID: 2ZCP). The two helices (indicated in orange) on which H12 and D27 are placed make up the ‘lid’ of the reaction cavity, and they are speculated to be fixated by the interaction with W38 (indicated in blue). (B) C₃₀ or C₄₀ synthase function of CrtM variants were determined as indicated in Fig. 4. Data points represent the averages of four replicates, and error bars represent the standard deviation.

Table 1. Mutations identified in CrtM variants.

Variant name	Mutation in addition to F26A
M1	A809 del (Frameshift)
M2	T398C (V133A), A718T (I240F), A726G
M3	T69C, T169G, A459G, A495G, G813 del (Frameshift)
M4	A561T (Q187H), A843 del (Frameshift)

Figure 1

E. coli
endogenous
MEP pathway

Figure 2

A

B

Figure 3

A

B

Figure 4

Figure 5

Figure 6

Figure 7