Biosynthesis of the Cyanobacterial Light-Harvesting Polypeptide Phyceroerythrocyanin Holo-α Subunit in a Heterologous Host

Aaron J. Tooley† and Alexander N. Glazer*

Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3200

Received 25 March 2002/Accepted 30 May 2002

The entire pathway for the biosynthesis of the phycobiliviolin-bearing His-tagged holo-α subunit of the cyanobacterial photosynthetic accessory protein phyceroerythrocyanin was reconstituted in Escherichia coli. Cyanobacterial genes encoding enzymes required for the conversion of heme to 3Z-phycocyanobilin, a precursor of phycobiliviolin (namely, heme oxygenase 1 and 3Z-phycocyanobilin:ferredoxin oxidoreductase), were expressed from a plasmid under the control of the hybrid trp-lac (trc) promoter. Genes for the apo-phyceroerythrocyanin α subunit (pecA) and the heterodimeric lyase/isomerase (pecE and pecF), which catalyzes both the covalent attachment of phycocyanobilin and its concurrent isomerization to phycobiliviolin, were expressed from the trc promoter on a second plasmid. Upon induction, recombinant E. coli used endogenous heme to produce holo-PecA with absorbance and fluorescence properties similar to those of the same protein produced in cyanobacteria. About two-thirds of the apo-PecA was converted to holo-PecA. No significant bilin addition took place in a similarly engineered E. coli strain that lacks pecE and pecF. By using immobilized metal affinity chromatography, both apo-PecA and holo-PecA were isolated as ternary complexes with PecE and PecF. The identities of all three components in the ternary complexes were established unambiguously by protein and tryptic peptide analyses performed by matrix-assisted laser desorption ionization–time-of-flight mass spectrometry.

Cyanobacteria produce macromolecular light-harvesting complexes made up largely of phycobiliproteins, called phycobilisomes. Each phycobiliprotein is an oligomer of a heterodimer, either a trimer ([αβ]3) or a hexamer ([αβ]6). Two of the major phycobiliproteins, allophycocyanin (λmax 650 nm) and phycocyanin (λmax ~620 nm), are found in all cyanobacterial phycobilisomes. Either phycocyanobilin (λmax 570 nm, with a shoulder at 595 nm) or phycocerythrin (λmax ~560 nm) may be present as well (12). Phyceroerythrocyanin occurs mainly in heterocystous cyanobacteria, and it has not been found together with phycocerythrin (4).

Different linear tetrapyrrole prosthetic groups (bilins), covalently attached through thioether linkages to cysteinyl residues, endow the phycobiliproteins with their distinctive light-harvesting properties. Phycocyanin carries one phycocyanobilin (PCB) on the α subunit (CpcA) at αCys84 and two PCBs on the β subunit at βCys82 and βCys155. Phyceroerythrocyanin carries one phycobiliviolin (PXB) on its α subunit (PecA) at αCys84 and, like phycocyanin, two PCBs on the β subunit (PecB) at βCys82 and βCys155 (3, 20). The first step in the pathway of bilin biosynthesis is the conversion of heme to biliverdin IXα (BVD) by the heme oxygenase encoded by hoxI (8). Recently, Frankenberg et al. (11) have shown that cyanobacterial pcyA genes encode bilin reductases that catalyze the four-electron reduction of BVD to 3Z-PCB. A limited amount of information is available on the genes that encode the proteins required for the covalent attachment of particular bilins at specific residues to apo-phycobiliprotein polypeptides.

Only two bilin lyases have been characterized: a heterodimeric phycocyanin α subunit PCB lyase encoded by the genes cpeE and cpeF (9, 10) and a phyceroerythrocyanin α subunit PCB lyase/isomerase encoded by the genes pecE and pecF (14, 21).

The structures of BVD, PCB, and PXB are shown in Fig. 1. Protein-linked PCB and PXB are isomers that differ as follows: there is a single bond between C-2 and C-3 in PCB and PXB and there is a double bond in PVB and PXB. In vitro studies by Zhao et al. (21) have shown that incubation of recombinant apo-PecA with PCB in the presence of PecE and PecF leads to the formation of holo-PecA and that PecE and PecF together catalyze both the addition of PCB to apo-PecA and its isomerization to PXB.

Recently, we reported the reconstitution of the pathway for the synthesis of holo-CpcA in Escherichia coli (19). In an E. coli strain, engineered with Synecocystis sp. strain PCC6803 genes, pcyA (encoding heme oxygenase 1) and pcyB (encoding 3Z-PCB:ferredoxin oxidoreductase) were expressed from one plasmid, and cpeA, cpeE, and cpeF were expressed from a second plasmid. The engineered E. coli strain expressed holo-CpcA with spectroscopic properties corresponding to those of the native CpcA produced in cyanobacteria. Here we describe the in vivo production of holo-PecA in an appropriately engineered strain of E. coli and compare the results with the outcomes of the in vitro experiments of Zhao et al. (21) and Storf et al. (18) for PCB addition to apo-PecA. Our studies also demonstrate the formation of a ternary complex of His-tagged PecA (HT-PecA) with PecE and PecF.
Anabaena sp. strain PCC7120 phycoerythrocyanin subunit. The primers used to amplify the pecA gene were 5'-GAG ATT AGG AGA CAT ATG AAA ACA CCT TTG ACC GAA GC-3' and 5'-CAA GAC CGA ATT CGA GTC TCT TAA CTT AAA GCG TTA ATT GCA TAG TTC AGG TA-3'. The resulting 0.5-kb product was digested with restriction enzymes SalI and NotI and cloned into SalI- and NotI-digested pBS350V, giving plasmid pBS435V. The primers used to amplify the pecE gene were 5'-ATT GTT CGG CCG AGG AGG AAC ATA TGA ATC AAG CCT CAT TGA GCG TAG ACG-3' and 5'-CCT GGA TCC GAG GCC GCT TTA AAG TGG TGT AAT TAA TTC TTT CTC CAG-3'. The resulting 0.55-kb product was digested with restriction enzymes SalI and NotI-digested pBS350V, giving plasmid pBS435V.

**Design and construction of the in vivo expression vector pBS437V containing the pecA-pecE-pecF cassette.** In order to engineer a strain of *E. coli* that could produce the phycoerythrocyanin holo-α subunit in vivo, an expression vector containing the genes involved in the production of holo-PecA (namely, ht-pecA, pecE, and pecF, where “ht” indicates the base sequence encoding the His tag) was designed. The aim was to introduce this plasmid, pBS437V, along with plasmid pAT101 (containing the Synechocystis sp. strain PCC6803 hox1-pecA cassette [19]) into *E. coli* and thereby produce all of the catalytic functions and components believed to be required for the formation of holo-PecA. The pecE gene, as a 0.8-kb SalI-NotI fragment from pBS435V, was cloned into SalI- and NotI-digested pBS430V containing ht-pecA, giving plasmid pBS436V. Subsequently, the pecF gene, as a 0.55-kb EagI-BamHI fragment from pBS434V, was cloned into EagI- and BamHI-digested pBS436V containing ht-pecA and pecE, giving the final cassette plasmid pBS437V (Fig. 2).

**Isolation of HT-PecA and associated proteins by immobilized metal affinity chromatography.** Cell pellets were thawed and resuspended in 20 ml of cold (0 to 4°C) buffer 0 (20 mM Tris-HCl [pH 8.0], 50 mM NaCl, 50 mM KCl).

Phenylmethylsulfonyl fluoride and 2-mercaptoethanol were added to final concentrations of 1 and 10 mM, respectively, immediately before breakage of cells by passage through a French pressure cell three times at 18,000 lbs/in². Cell debris was removed by centrifugation at 4°C in a Beckman JA20 rotor at 30,000 × g for 1 h. The supernatant solution was mixed with 2 to 3 ml of Ni²⁺-NTA agarose at 4°C for 15 min before the agarose was loaded onto a column. The agarose was then washed with 10 column volumes each of cold buffer A1 (20 mM Tris-HCl [pH 8.0], 50 mM NaCl, 50 mM KCl, 20 mM imidazole, 5% [vol/vol] glycerol), buffer B (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 500 mM KCl), and
RESULTS

Construction of *E. coli* expression strains. In an earlier study, we introduced into an *E. coli* strain the gene encoding a His-tagged version of the apo-α subunit of phycoerythrin (*pecA*), as well as the genes encoding the components of the biosynthetic pathway in cyanobacteria leading from heme to PCB (*hoc1* and *pcy4*) and the genes encoding the lyase (*pecE* and *pecF*) required for the formation of the site-specific cysteinyl adduct of PCB with CpcA. The engineered *E. coli* produced holo-CpcA with spectroscopic properties characteristic of the native protein (19).

The genes (*pecE* and *pecF*) encoding the heterodimeric phycoerythrin α subunit PCB lyase have been characterized (14). Moreover, Zhao et al. (21) and Storf et al. (18) have shown that in vitro addition of PCB to a mixture of recombinant HT-PecA, PecE, and PecF leads to the production of holo-PecA carrying a PXB chromophore. Consequently, PecE and PecF must a catalyze reaction(s) in which PCB is both attached to PecA and isomerized to PXB, although it is not clear whether the isomerization involves the thioether-bound or free PCB.

The information described above leads to the prediction that an *E. coli* strain engineered to contain *hoc1*, *pcy4*, *ht-pecA*, *pecE*, and *pecF* should produce holo-HT-PecA with spectroscopic properties characteristic of the native holo subunit by the pathway illustrated in Fig. 1. The His tag on PecA is not expected to cause any difficulties. HT-PecA expressed in *Anabaena* sp. strain PCC7120 has native spectroscopic properties and is recovered by immobilized metal affinity chromatography as a 1:1 complex with holo-PxB that is efficiently incorporated into the phycobilisome (Y. A. Cai and A. N. Glazer, unpublished data).

To test the prediction described above, two strains of *E. coli*, each carrying two expression vectors, were generated. To maintain both plasmids simultaneously, double transformants were selected for resistance to both spectinomycin and kanamycin and were grown with both spectinomycin and kanamycin. *E. coli* strain DH5α(347V,101) carries plasmids pBS437V (ht-pecA–pecE–pecF) cassette, Kmr) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2). Another *E. coli* strain, DH5α(343V,101), carries plasmid pBS437V (ht-pecA–pecE–pecF) cassette, Sp<sup>+</sup>) and pAT101 (ht-hoc1–pcy4A cassette, Km<sup>+</sup>) and expresses all the known genes believed to be involved in PCB biosynthesis and in addition isomerizes PCB to apo-PecA (Fig. 2).
PecE and PecF for PCB addition and isomerization to PXB in vivo.

Heterologous expression and characterization of holo-HT-PecA. Expression of the gene cassettes in the two strains of E. coli led to two very different phenotypes. Upon induction with IPTG, the E. coli strain DH5α(437V) culture acquired a pronounced pink tint. This color change was not seen in the culture of DH5α(430V,101). HT-PecA was purified from strain DH5α(437V,101) by affinity chromatography. Analysis of Coomassie brilliant blue-stained gels after SDS-PAGE of the purified protein showed that there were three distinct bands (Fig. 3, lane 2, upper panel), band a at ~26 kDa, band b at ~20 kDa, and band c at ~17.5 kDa. The sizes of these bands are compatible with those calculated for PecE, holo-HT-PecA, and PecF, respectively (see below). Upon exposure to Zn²⁺ (1) and UV illumination, only band b was fluorescent, indicating that it contained covalently attached bilin (Fig. 3, lane 2, lower panel). The mobility of band b was the same as that of band d, which is holo-HT-PecA expressed in Anabaena sp. strain PCC7120, and HT-HO1 was not seen in the gel due to its low expression levels.

The spectroscopic properties of holo-HT-PecA were determined with the protein fraction purified by immobilized metal affinity chromatography after dialysis against 50 mM Na phosphate (pH 7.0). The absorbance spectrum of holo-HT-PecA had a λ<sub>max</sub> at 561 nm and a λ<sub>ex</sub> at 582 nm (Fig. 4). The spectrum corresponded closely to that of the native holo-PecA obtained from Anabaena sp. strain PCC7120 (5).

Storf et al. (18) reported an ε<sub>577</sub> of 77,000 M⁻¹cm⁻¹ at 567 nm for holo-HT-PecA produced in vitro by the addition of PCB to a mixture of HT-PecA, PecE, and PecF in 100 mM phosphate buffer at pH 7.0. By using this ε<sub>577</sub> and a molecular weight of 20,920 for holo-HT-PecA, it can be calculated that ~0.05 mg of purified holo-HT-PecA was recovered per g (wet weight) of E. coli cells under the expression conditions described here.

Very little soluble HT-PecA was recovered upon cell breakage of induced E. coli strain DH5α(430V,101) (ht-pcA, h-trox1–pcyA). A similar result was seen with E. coli strain DH5α(430V), in which only HT-PecA is expressed. However, when the ht-pcA–pecE–pecF gene cassette was expressed in E. coli from plasmid pBS437V, apo-HT-PecA was purified from the soluble fraction by affinity chromatography with a yield similar to that seen with E. coli strain DH5α(437V,101).

E. coli strains producing PecA, PecE, and PecF show approximately one-third slower growth than E. coli strains expressing similar, but different, phycobiliprotein cassettes (data not shown).

Characterization of a holo-HT-PecA–PecE–PecF ternary complex. As described above, analysis of affinity-purified holo-HT-PecA produced in E. coli strain DH5α(437V,101) by SDS-PAGE showed the presence of three protein components whose sizes were compatible with those calculated for PecE, holo-HT-PecA, and PecF (Fig. 3, lane 2). Analysis of affinity-purified HT-PecA produced in E. coli strain DH5α(437V) likewise showed the presence of three protein components with the same electrophoretic mobilities on SDS-PAGE gels, except that the HT-PecA did not carry bilin (Fig. 3, lane 1).

The protein fractions purified from the two E. coli strains by immobilized metal affinity chromatography were subjected to analysis by MALDI-TOF mass spectrometry. For the fraction

FIG. 3. SDS-PAGE analysis of Anabaena sp. strain PCC7120 His-tagged proteins expressed in various E. coli strains and purified by immobilized metal affinity chromatography. Proteins were visualized by Coomassie brilliant blue staining (upper panel) and by the UV-excited fluorescence of bilin-bearing polypeptides in the presence of Zn²⁺ (1) (lower panel). Lane 1, PecE, HT-PecA, and PecF (bands a, b, and c, respectively) expressed in E. coli strain DH5α(437V); lane 2, PecE, HT-PecA, and PecF expressed in strain DH5α(437V,101), which also expresses HT-HO1 (band not in gel) and PcyA; lane 3, molecular mass standards (from the top, 30, 25, 20, and 15 kDa); lane 4, holo-HT-PecA (band d) and PecB (band e) from Anabaena sp. strain PCC7120 (see text).

FIG. 4. Spectroscopic properties of Anabaena sp. strain PCC7120 holo-HT-PecA expressed in E. coli. Shown are the absorbance (λ<sub>max</sub>, 561 nm) (thick solid line), control Anabaena sp. strain PCC7120 holo-PecA (from the study of Bryant et al. [5]) (thin solid line), and fluorescence emission (λ<sub>exc</sub>, 582 nm; λ<sub>em</sub>, 550 nm) (dashed line) spectra for HT-PecA purified from E. coli by affinity chromatography. The spectra were normalized at 561 nm to the absorbance spectrum of holo-PecA from E. coli.
HT-PecA

| MGRHSHHHHDD GTPSPLTF TGQAKNSYPLG EAISSADVVG STLSNTEMQA VFGYFWRARA | 60 |
| GLAAQAFSN GKFNAEEAA NNYYQEPYFT TCQSGQYAS TEPGKSKCM DNDMYGTSI | 120 |
| YCPPVTVGSF LEDYVVSGLS ELNSALGSLP SWYVAALEFV EIDHGLGDUV AGEAIYNLH | 180 |
| AINALS | 186 |

PecF

| MNQALSLSYDA ITNLIEAFHH HHPAVRSAV DELIKGISST VNLILAYDD SQDQGQAOQT | 60 |
| IQVLAQIQDA KALELLAEVV GTSVANCGG NVSRBAARGL GEIATSTGTN EIINIAQELK | 120 |
| IWALLRPEDW GLRYAAAVSL QEIATTPKRA ALQQAAMQRT DPVVRSMAT ALS | 173 |

FIG. 5. Amino acid sequences of HT-PecA, PecE, and PecF, showing in boldface type the residues in peptides identified by mass spectrometry of tryptic digests of proteins in gel slices from bands b, a, and c, respectively, cut out from gels such as the gel shown in Fig. 3, lanes 1 and 2.

The studies described here documented successful reconstitution of the predicted pathway for the biosynthesis of the phycoerythrocyanin holoprotein. The results of our experiments also support the conclusion reached by Zhao et al. (21) and Storf et al. (18) on the basis of in vitro experiments that the heterodimeric lyase/isomerase (PecE/PecF) catalyzes both the covalent attachment of PCB and the concurrent isomerization of the molecule to PXB.

Storf et al. (18) also reported in vivo nonenzymatic formation of a PCB adduct with apo-HT-PecA. We saw no addition of PCB to HT-PecA in E. coli in the absence of PecE and PecF. A caveat is that little of the PecA is in a soluble form in the absence of PecE and PecF. Storf et al. (18) noted that, "The overexpressed proteins are deposited mainly in inclusion bodies, but can be well solubilized by extended sonication." It is possible that the protein solubilized by sonication is in an altered conformation that allows PCB addition.

We show here that the presence of PecE and PecF markedly increases the solubility of apo-HT-PecA and that the three proteins form a complex that can be isolated by affinity chromatography. A similar result is seen with holo-HT-PecA (Fig. 3, lanes 1 and 2). In contrast, when holo-HT-CpcA was purified from E. coli that also expressed CpcE and CpcF, the latter did not copurify with holo-HT-CpcA during affinity chromatography (19). PecF has four consecutive His residues at positions 19 to 22 (Fig. 5), and it is possible that added interaction of these residues with Ni²⁺-NTA agarose may contribute to the stability of the HT-PecA–PecE–PecF complex during chromatographic purification.

Jung et al. (14) established that in Anabaena sp. strain PCC7120 interposon mutagenesis of pecE, pecF, or pecE and pecF eliminates the formation of holo-PecA. They showed that ~30% of the wild-type level of PecB was present in the phycobilisomes of all of the mutants. In contrast, holo-PecA was barely detectable in the pecE and pecF mutants, but it was present in the pecEF deletion mutant as a PCB-adduct in a 1:1 ratio with the PecB. Jung et al. confirmed the identity of this unnatural adduct by isolation of the subunit and amino-terminal sequencing.

Jung et al. (14) commented on these results as follows: "A testable hypothesis is that when the cognate lyase is missing, the phycoerythrosin α subunit lyase ‘rescues’ the phycoerythrocyanin α subunit by adding PCB. How is the absence of such α
PCB adduct of PecA in the PecE and PecF mutants explained in the context of this hypothesis? When only PecE or PecF is present either of these components (or a hybrid PecE/CpcF or PecF/CpeE lyase) may form a complex with the Pec apo-α subunit which is thereby made unavailable to the CpeEF lyase.” Our demonstration of the formation of a PecA-PecE-PecF complex provides support for this speculation.

Holo-HT-PecA expressed in Anabena sp. strain PCC7120 and purified by immobilized metal affinity chromatography is recovered as a 1:1 complex with holo-PecB (Cai and Glazer, unpublished results). It is thus possible that in cyanobacteria release of holo-HT-PecA from the PecA-PecE-PecF complex is mediated by the formation of the holo-PecE–holo-PecB heterodimer.

The ability to express in heterologous hosts phycobiliprotein subunits with distinctive spectroscopic properties, such as holo-PecA (λmax = 561 nm; λF = 582 nm) and holo-CpeA (λmax = 625 nm; λF = 641 nm) (19), provides the potential for new fluorescent labels, which are valuable in cell biology, whose spectroscopic properties extend the range of wavelengths accessible to the widely used green fluorescent protein fusions (7, 13) and to the Discosoma red fluorescent protein (2).

ACKNOWLEDGMENTS

We are grateful to Arnold Falick and Sharleen Zhou of the HHMI Mass Spectrometry Laboratory of the Department of Molecular and Cell Biology at the University of California, Berkeley, for performing the mass spectrometric analyses. We thank Yuping A. Cai for helpful discussions and Cynthia Voong for her assistance.

This work was supported in part by a grant from the Lucille P. Markey Charitable Trust.

REFERENCES