De novo biosynthesis of a nonnatural cobalt porphyrin cofactor in E. coli and incorporation into hemoproteins

Lydia J. Perkinsa,1, Brian R. Weavera,1, Andrew R. Bullera,2, and Judith N. Burstyna,2,3

*Department of Chemistry, University of Wisconsin–Madison, Madison, WI 53706

Edited by William F. DeGrado, University of California, San Francisco, CA, and approved March 1, 2021 (received for review August 19, 2020)

Enzymes that bear a nonnative or artificially introduced metal center can engender novel reactivity and enable new spectroscopic and structural studies. In the case of metal-organic cofactors, such as metalloporphyrins, no general methods exist to build and incorporate new-to-nature cofactor analogs in vivo. We report here that a common laboratory strain, Escherichia coli BL21(DE3), biosynthesizes cobalt protoporphyrin IX (CoPPIX) under iron-limited, cobalt-rich growth conditions. In supplemented minimal media containing CoCl2, the metabolically produced CoPPIX is directly incorporated into multiple hemoproteins in place of native heme b (FePPIX). Five cobalt-substituted proteins were successfully expressed with this new-to-nature cobalt porphyrin cofactor: myoglobin H64V V68A, dye decolorizing peroxidase, aldoxime dehydratase, cytochrome P450 119, and catalase. We show conclusively that these proteins incorporate CoPPIX, with the CoPPIX making up at least 95% of the total porphyrin content. In cases in which the native metal ligand is a sulfur or nitrogen, spectroscopic parameters are consistent with retention of native metal ligands. This method is an improvement on previous approaches with respect to both yield and ease-of-implementation. Significantly, this method overcomes a long-standing challenge to incorporate nonnatural cofactors through de novo biosynthesis. By utilizing a ubiquitous laboratory strain, this process will facilitate spectroscopic studies and the development of enzymes for CoPPIX-mediated biocatalysis.

Significance

Enzymes rely on cofactors, both organic molecules and metal ions, for reactivity beyond what can be performed with the 20 standard amino acids. One way to expand the reactivity of heme-dependent proteins is to swap the native iron atom in the heme cofactor for a different metal. Because the resulting metal-containing cofactor is new to nature, methods for producing these nonnative hemeproteins are difficult and often limited to the test tube. Here, we report a method for subverting the heme biosynthetic pathway in E. coli to produce cobalt protoporphyrin IX (CoPPIX) and CoPPIX-substituted hemoproteins in living cells. This method enables the use of CoPPIX-substituted enzymes as biocatalysts and widens the scope of proteins that may be studied by metal substitution.

Author contributions: L.J.P. and B.R.W. performed research; L.J.P., B.R.W., A.R.B., and J.N.B. designed research, analyzed data, and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

1L.J.P. and B.R.W. contributed equally to this work.
2To whom correspondence should be addressed. Email: arbuller@wisc.edu or burstyn@chem.wisc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2017625118/-/DCSupplemental.

Published April 13, 2021.

PNAS 2021 Vol. 118 No. 16 e2017625118
https://doi.org/10.1073/pnas.2017625118 | 1 of 10
make metal-substituted hemoproteins (41–46). Early on, copper, cobalt, nickel, and manganese-substituted horseradish peroxidase (HRP) were prepared by a multistep process that subjected protein to strong acid and organic solvents (41, 42). Variations of this method have been used repeatedly (24, 43, 47–49). However, this method is applicable only to a narrow range of hemoproteins that tolerate the harsh treatment. With the advent of overexpression methods, significant improvement of metalloporphyrin-substituted protein yield was achieved by direct expression of the apoprotein and reconstitution with the desired metalloporphyrin in lysate prior to purification (50). Although this approach has many virtues, direct expression of apoprotein is ineffective for many hemoproteins, again limiting the utility of this method.

As an alternative to the above in vitro approaches, researchers have pursued systems for direct in vivo expression of metal substituted hemoproteins. Two specialty strains of *Escherichia coli* (*E. coli*) were engineered to incorporate metalloporphyrin analogs from the growth medium into hemoproteins during protein expression. The engineered RP523 strain cannot biosynthesize heme and bears an uncharacterized heme permeability phenotype. Together, these two features enable this strain to assimilate and incorporate various metalloporphyrins into overexpressed hemoproteins with no background heme incorporation (44, 51–53). However, heme auxotrophy makes RP523 cells exceedingly sensitive to O2, and, in many situations, RP523 cultures must be grown anaerobically. An alternative BL21(DE3)-based engineered strain harbors a plasmid bearing the heme transporter ChuA, which facilitates import of exogenous heme analogs (45). Production of metalloporphyrin-substituted protein with this ChuA-containing strain relies on growth in iron-limited minimal media, thereby diminishing heme biosynthesis. This method was used successfully to express metal-substituted versions of the heme domain of cytochrome P450 BM3 (45) and several myoglobin variants (11, 12). Because these cells biosynthesize a small quantity of their own heme, they are far more robust than the RP523 cells. Unfortunately, this advantage comes at the cost of increased heme contamination in the protein product (2 to 5%) (45).

A set of intriguing papers reported the production of cobalt-substituted human cystathionine β-synthase (CoCBS) that relies on the de novo biosynthesis of CoPPIX from CoCl2 and δ-aminolevulinic acid (δALA), a biosynthetic precursor to heme (46, 54). This method yielded significant amounts of CoCBS—albeit with modest heme contamination (7.4%)—sufficient for spectroscopic and functional characterization of the CoPPIX-substituted protein (8, 46). As cobalt is known to be toxic to *E. coli*, the researchers passaged the CBS expression strain through cobalt-containing minimal media for 12 d, enabling the cells to adapt to high concentrations of cobalt prior to protein expression. It is plausible that this serial passaging alters the *E. coli* cells, enabling the biosynthesis of CoPPIX and in vivo production of metal-substituted protein. The adaptation process is slow (>10 d), and it is unknown how genomic instability under these mutagenic conditions affects the reproducibility of this passaging approach.

The possibility of facile CoPPIX production is particularly attractive for future biocatalysis efforts. As described above, synthetic cobalt porphyrins have been shown to perform a range of radical-mediated reactions. The ability to produce a CoPPIX center in vivo may enable engineering these unusual reactivities via directed evolution in addition to spectroscopic applications. We therefore set out to explore the unusual phenotype of CoPPIX production by *E. coli* and to ascertain whether it was possible to efficiently biosynthesize cobalt-containing hemoproteins in vivo from a single “generalist” cell line. Our goal was to achieve an efficient and facile method of cobalt-substituted hemoprotein production with minimal contamination of the native cofactor. Herein, we report the surprising discovery that native *E. coli* BL21(DE3) can biosynthesize a new-to-nature CoPPIX cofactor (Fig. 1). We use this insight to produce cobalt-substituted hemoproteins in vivo without requirement for complex expression methods or specialized strains.

**Results and Discussion**

**Cobalt-Tolerant *E. coli* Are Not Optimal for CoPPIX-Substituted Hemoprotein Expression.** We envisioned a strategy for producing CoPPIX-substituted hemoproteins that would utilize a simple cobalt tolerant mutant strain as the source of the metalloporphyrin. This approach was inspired by the procedure reported by Majtan et al. (46), in which a CBS-expressing strain was serially passaged until direct expression of CoCBS was achieved. We hypothesized that this serial passaging method might select for mutant strains with a cobalt tolerance phenotype, enabling CoPPIX production. Therefore, we attempted to evolve a cobalt-tolerant strain that could be used as a “generalist” to express any CoPPIX-substituted hemoprotein. To this end, we passaged *E. coli* BL21(DE3)—a commonly used laboratory expression strain—through cobalt-supplemented media to select for cobalt-tolerant phenotypes. Specifically, we subjected *E. coli* to 15 rounds of passaging over 12 d (*SI Appendix, Fig. S1*). To initiate each passage, the previous culture was diluted into fresh CoCl2-containing, augmented minimal media (M9*) that was supplemented with low concentrations of δALA to facilitate porphyrin production. The concentration of CoCl2 in the media was increased in a stepwise fashion after the cells achieved a steady, high growth rate and a stable cell density at stationary phase. The resultant strains were able to grow in much higher concentrations of cobalt than the parent strain, with an IC50 (inhibitory concentration, i.e., the concentration at which growth is inhibited by 50%) that was 10-fold higher than that of the parental strain of *E. coli* (*SI Appendix, Table S1 and Fig. S5*). Further experimental details may be found in *SI Appendix*.

To determine whether this improved cobalt tolerance enabled production of a CoPPIX-substituted hemoprotein, we tested expression of a sperm whale Mb variant H64V V68A bearing a 6-His tag on the C terminus (Mb*). This Mb variant, in which oxygen binding is eliminated by substitution of His64, was engineered for cyclopropanation of styrene (55) and recently was used by Sreeniyay et al. to carry out a variety of carbene transfer reactions using CoPPIX and other noniron-containing metalloporphyrins (11). Importantly, electronic absorption spectra of the CoPPIX-substituted Mb* were reported, and the cobalt-substituted native Mb protein has been well characterized (3–5, 56, 57). Mb* was overexpressed in parent BL21(DE3) cells and cells from one passaged strain (strain A15, *SI Appendix*) in augmented minimal media (M9*) supplemented with 500 μM CoCl2 and 250 μM δALA. The resulting Mb* proteins were purified by metal affinity chromatography. To compare the effect of passaging on the metalloporphyrin incorporation, the two purified Mb* samples were analyzed for the presence of CoPPIX and FePPIX by basic extraction of the pyridine (py)-coordinated metalloporphyrin complexes (58). The Co(II)PPIX(py) and Fe(II)PPIX(py)3 complexes exhibit distinct visible spectra, which can be used for metalloporphyrin identification (46). Expression in parent and passaged strains resulted in an appreciable amount of CoPPIX-substituted Mb* (*SI Appendix, Fig. S2*), with protein yields of 3 to 5 mg·L−1 culture.

To our surprise and disappointment, the Mb* purified from the cobalt-resistant strain contained significant amounts of FePPIX. In contrast with the results from Majtan et al. (46), in which the

![Fig. 1. Chemical structures of iron protoporphyrin IX (FePPIX or heme b), cobalt protoporphyrin IX (CoPPIX), and free base protoporphyrin IX (H2PPIX).](https://doi.org/10.1073/pnas.2017625118)
fraction of native heme in the isolated CoCBS was less than 10%, in Mb*, the visible spectrum of the pyridine-extracted cofactor suggests that close to half of the incorporated metalloporphyrin cofactor was heme (SI Appendix, Fig. S2). This observation was unexpected given the iron limited growth conditions. Inductively coupled plasma mass spectrometry (ICP-MS) analysis of the iron content in the M9* expression medium revealed less than 1 ppb Fe. We hypothesize that the cobalt-tolerant strain has an increased capacity to assimilate iron and to discriminate iron from cobalt through unknown mechanisms. While passing the BL21(DE3) cells, we deliberately increased the cobalt concentration as the cells adapted, with the result that the passaged cells grew significantly better in high cobalt (1 mM CoCl₂) media. This cobalt concentration was significantly higher than that employed by Majtan et al. (150 μM) (46, 54). It is of note that in the prior CoCBS expression, the passaged cells contained a plasmid encoding either CBS (a heme-containing protein) or an FeS cluster–containing protein. In our experiments, no plasmid was present during the passaging process. Whether these differences altered the outcome of the passaging process and its previously reported ability to cleanly produce CoPPIX-containing protein is unknown.

Remarkably, the Mb* sample expressed in the parent BL21(DE3), which was intended as a negative control, contained high levels of CoPPIX with no apparent FePPIX contamination. The visible spectrum of the pyridine-extracted cofactor from isolated Mb* is identical to that of the CoPPIX standard, with no evidence of heme (SI Appendix, Fig. S2). The efficient production of CoMb* was completely unexpected, especially since the cells were grown at a toxic cobalt concentration (500 μM). The measured IC₅₀ for CoCl₂ in E. coli BL21(DE3) was 22 μM, (SI Appendix, Fig. S5), yet cells expressing Mb* grew effectively and produced good yields of CoMb*, motivating further investigation of these observations.

**E. coli** BL21(DE3) Can Biosynthesize CoPPIX De Novo in Cobalt-Rich, Iron-Poor Conditions. To characterize the effects of cobalt on E. coli growth and porphyrin production, we measured the ability of E. coli BL21(DE3) to grow in the presence and absence of high levels of iron and cobalt (500 μM), while supplementing with exogenous δALA to drive heme biosynthesis. Growth curves comparing the effects of these additives to M9* media on E. coli growth rate are shown in Fig. 2B. Cells were harvested from each culture when growth reached stationary phase, and porphyrins were extracted with acidified pyridine and analyzed via ultra-performance liquid chromatography mass spectrometry (UPLC-MS) (Fig. 2B). E. coli BL21(DE3) cells grew well in M9* broth and produced relatively little porphyrin. Cell cultures supplemented with δALA and no extra metal grew more slowly and produced a large quantity of free-base protoporphyrin IX (H₂PPIX). Cell cultures coproduced with δALA, and FeCl₃ showed a dramatic increase in growth rate to significantly higher cell density. These cells produced substantial FePPIX, with a small amount of residual H₃PPIX. Cultures cosupplemented with δALA and CoCl₂ showed a dramatic lag in growth of almost 24 h and grew to a lower cell density in stationary phase. This result is likely a consequence of cobalt-induced disruption of iron metabolism, which is known to cause defects in iron–sulfur cluster biosynthesis and induce oxidative stress (59, 60). Cell pellets isolated from these cobalt-supplemented cultures were distinctly pink, suggesting the presence of a different porphyrin species (SI Appendix, Fig. S3). Indeed, extraction and analysis revealed significant CoPPIX content in these cells. Free-base H₃PPIX was undetectable, and the amount of FePPIX produced was comparable to that of the δALA-supplemented sample that lacked additional FeCl₃. These data clearly show that, in the presence of near-toxic levels of CoCl₂ and without any passaging in cobalt-rich media, E. coli BL21(DE3) cells produce significant amounts of CoPPIX.

This observation raises questions as to the origin of this nonnative metalloporphyrin cofactor. The final step of heme biosynthesis is divalent iron insertion into H₃PPIX, catalyzed by the enzyme ferrochelatase. Notably, the ferrochelatase homologs from *Homo sapiens*, *Mus musculus*, *Saccharomyces cerevisiae*, and *Bacillus subtilis* have exhibited promiscuous activity in vitro with the alternative divalent metal ions cobalt and zinc (48–51). Although the metal specificity of *E. coli* ferrochelatase has not been determined, the data reported here suggest that this enzyme may possess promiscuous activity with cobalt, resulting in the biosynthesis of CoPPIX when E. coli cells are grown in the presence of high concentrations of cobalt and minimal iron. Given the ability of *E. coli* to biosynthesize CoPPIX, we sought to explore the properties of BL21(DE3) as an expression strain for cobalt-substituted hemoproteins.

**Fig. 2.** The effect of iron, cobalt, and δALA on E. coli BL21(DE3) growth and porphyrin production. (A) OD₆₀₀ values monitoring cell growth over time in M9* media without additives (black) and supplemented with 250 μM δALA (yellow), 500 μM FeCl₃ and 250 μM δALA (red), and 500 μM CoCl₂ and 250 μM δALA (blue). Error bars are the SD among OD₆₀₀ values from three separate cultures. (B) Porphyrin content in cells harvested in stationary phase from A as measured by UPLC-MS. Amount of porphyrin calculated is shown relative to the OD₆₀₀ of cells when collected. See SI Appendix, General Experimental Methods for details.
ferred greater tolerance to CoCl2 than expression of tested whether expression of heme-containing protein Mb* con-
tryptophan synthase from expressed at high levels, we chose the PLP-containing β-subunit of tryptophan synthase from Pyrococcus furiosus (PYTrpB) (61). We tested whether expression of heme-containing protein Mb* conferred greater tolerance to CoCl2 than expression of PYTrpB. To measure the IC50 of CoCl2 on E. coli growth, cells containing isopropyl β-D-1-thiapactryptopyranoside (IPTG)-inducible, protein-
coding plasmids were grown in 96-well plates with variable concentrations of IPTG and CoCl2. Growth was measured via OD600 using a plate reader (SI Appendix, Fig. S4). Overexpression of PYTrpB in unsupplemented M9* broth results in a small decrease in cell density relative to nonoverexpression (no IPTG added) controls, consistent with diversion of resources to protein production (Fig. 3A). Nevertheless, the IC50 for CoCl2 was not significantly altered upon overexpression of the PYTrpB protein. As was observed upon expression of PYTrpB, expression of Mb* in unsupplemented M9* broth results in a small decrease in cell density. In contrast to the PYTrpB strain, overexpression of Mb* also causes a significant increase in the IC50 for CoCl2 (Fig. 3 and SI Appendix, Fig. S7), revealing that heme protein overexpression indeed helps to alleviate cobalt toxicity. Therefore, we adopted an approach to expressing CoPPIX-substituted hemoproteins utilizing relatively high concentrations of CoCl2 (500 μM), in which heme protein overexpression exerts a selective advantage on cell growth.

**E. coli BL21(DE3) Is an Effective Host for In Vivo Production of CoMb**

We next established a method to produce CoMb* on a preparative scale using a two-step induction procedure. E. coli BL21(DE3) cells were grown in M9* broth (1.0 LiNa, 2.8 Fe, Bac flask) to an intermediate OD600 of 0.2 to 0.3, whereupon δALA (250 μM) and CoCl2 (500 μM) were added to stimulate formation of CoPPIX. Cells were grown in this supplemented medium to an OD600 of 0.6 to 0.8, after which Mb* expression was induced by addition of IPTG (250 μM). The overexpressed Mb* protein was purified using Ni-affinity chromatography and stored at −80 °C.

The logic of this expression protocol, in which the cells were grown in unsupplemented M9* to a moderate cell density and then with δALA and CoCl2, prior to induction of Mb* expression, is as follows. We observed that cells inoculated at low density in cobalt-containing media exhibit a substantial growth lag, which may be offset by allowing the cells to grow in cobalt-free media. Further, we were aware that the medium contained trace iron.

![Fig. 3. Effect of overexpressing a heme protein or a nonheme protein on cobalt tolerance of E. coli BL21(DE3). OD600 measured after 9 h of growth in variable amounts of CoCl2.](https://www.pnas.org/content/10.703/pnas.2017625118)

Given the high efficiency with which iron is incorporated into heme, the preincubation period was designed to allow full incorporation of all available heme into native hemoproteins. In the next phase, addition of δALA and CoCl2, CoPPIX is produced. Under these conditions, we anticipated that CoPPIX was present and available for incorporation into Mb* at the time of induction. While this method provides a pool of CoPPIX, residual FePPIX or H2PPIX might also be available for incorporation into hemoproteins. We therefore assessed the identity of the porphyrin cofactor bound within the isolated Mb* using three independent methods: pyridine extraction assay, ICP-MS, and UPLC-MS. First, we extracted the porphyrin cofactors from purified Mb* and characterized the metalloporphyrin pyridine complexes using visible spectroscopy. The spectral peaks positions and band shapes for the extracted porphyrins closely matched that of a Co(II)PPIX(py) standard (SI Appendix, Fig. S8A). The absence of any shoulder at 424 nm, the Soret peak position of Fe(II)PPIX(py)2, suggested that there was minimal FePPIX in the Mb* protein. To quantify the metal content of Mb* more rigorously, we employed ICP-MS. From ICP-MS analysis, we determined that the vast majority of the metal content (>99.5%) was cobalt, and less than 0.5% was iron (Table 1). We further considered the possibility that the total metalloporphyrin content and metal content may not fully reflect the entire porphyrin content in the protein sample, as some porphyrin may be unmetallated. Therefore, we assessed the relative amounts of H2PPIX, FePPIX, and CoPPIX by UPLC-MS. UPLC-MS analysis of isolated Mb* confirmed that the vast majority of the porphyrin present in the protein was CoPPIX, representing greater than 99% of the total porphyrin. FePPIX and H2PPIX were present at less than 1% total porphyrin. These three experiments together revealed that the isolated Mb* contained CoPPIX essentially to the exclusion of other porphyrin cofactors.

Because heterologous overexpression may result in significant apoprotein, we assessed the fraction of isolated Mb* that was loaded with CoPPIX. The total cobalt concentration was measured by ICP-MS and compared to the protein concentration, as measured by bicinchoninic acid (BCA) assay. According to the measured CoPPIX to total Mb* ratio (Table 1), the isolated Mb* protein is only 30% loaded with CoPPIX. This observation suggests two nonexclusive possibilities. First, Mb* production outpaces CoPPIX production under the expression conditions. Second, CoPPIX may not bind efficiently to Mb*. Substoichiometric cofactor loading is frequently observed for hemoproteins and has previously been ascribed to incommensurate rates of apoprotein production and heme biosynthesis under overexpression conditions (62, 63).

We confirmed that native histidine coordination (Fig. 4A) was maintained in CoMb* using electronic absorption and EPR spectroscopy. Electronic absorption spectra of the as-isolated Co(II)Mb* and dithionite-reduced Co(II)Mb* (Fig. 4B) are similar to those of...
CoMb* and CoMb previously produced by two different routes: in vivo expression of Mb* in engineered E. coli cells capable of importing CoPPIX from the media (11) and by reconstitution of apoMb with CoPPIX (24) (Table 2). In the Co(III)Mb* spectrum, we observe a major set of peaks, listed in Table 2, and a minor set of peaks, the latter appearing at 404, 520, and 561 nm. The presence of two sets of peaks clearly implicates the presence of two differently coordinated species. The major species exhibits peak maxima comparable to those of Co(III)Mb, in which cobalt is six-coordinate—bound by His94 and a water molecule (Table 2). The identity of the minor species is less clear. The distal pocket of this Mb* variant is hydrophobic and lacks the usual hydrogen bond—donating His64. We suggest the minor species is a five-coordinate complex lacking the sixth water ligand based on analysis of studies of other cobalt porphyrin complexes (64, 65). The relatively slow reduction of Co(III)Mb*, which requires 30 min to complete, is consistent with previous observations of the behavior of CoMb (56).

To conclusively establish the presence of the native histidine ligand in CoMb*, we turned to EPR spectroscopy. The EPR spectrum of CoMb* (Fig. 4C) matches those previously reported for CoMb prepared by other methods (4). The presence of triplet superhyperfine coupling unequivocally confirms that cobalt is coordinated by a nitrogen donor ligand. As previously observed for CoMb, the Co(II)Mb* EPR spectrum is axially symmetric, consistent with a low-spin Co(II) complex with an unpaired electron. The octaplet hyperfine and triplet superhyperfine structures are indicative of coupling to S = 7/2 Co and S = 1/2 N nuclei, respectively (Fig. 4C). Spin Hamiltonian parameters derived from simulation (Table 3) match well with those previously reported for wild-type sperm whale deoxy-Co(II)Mb (4), in which the nitrogen superhyperfine coupling is attributed to axial nitrogen coordination by His94. These spectroscopic data are wholly consistent with a natively bound CoPPIX cofactor in Mb*.

Encouraged by our success with expression of CoMb*, we considered to what extent other CoPPIX-substituted hemoproteins might be produced by this method. We chose myoglobin as a starting point because it is an excellent model protein on which to test CoPPIX incorporation: apoMb is relatively stable and easy to express, heme has previously been introduced by multiple routes, and spectroscopic data on CoMb are available for comparison. While the stability and versatile cofactor scavenging properties of apoMb are beneficial for incorporating nonnative porphyrins, its stability may be responsible for the large proportion of apoMb that we observed in isolated CoMb*. This stability of the apo form of the protein, paired with the solvent-accessible heme-binding pocket, may facilitate in vitro reconstitution. In contrast, in vivo heme insertion into myoglobin is hypothesized to be cotranslational (66, 67) and may also be aided by chaperones (68, 69). A more stringent challenge of our method of direct in vivo expression is production of diverse hemoproteins, including those for which in vitro reconstitution is inefficient or untested.

**Table 1. Comparison of the relative porphyrin content, relative metal content, cofactor loading, and protein yield for CoPPIX-substituted hemoproteins produced by overexpression in CoCl2-supplemented media**

<table>
<thead>
<tr>
<th>Protein</th>
<th>CoPPIX</th>
<th>FePPIX</th>
<th>H2PPIX</th>
<th>Co</th>
<th>Fe</th>
<th>% holoprotein</th>
<th>Yield (mg · L⁻¹ culture)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoMb H64V V68A (Mb*)</td>
<td>&gt;99%</td>
<td>0.1 ± 0.1%</td>
<td>0.1 ± 0.1%</td>
<td>99.8 ± 0.3%</td>
<td>0.2 ± 0.3%</td>
<td>29 ± 2</td>
<td>4.7</td>
</tr>
<tr>
<td>CoCPY119</td>
<td>99 ± 1%</td>
<td>&lt;1%</td>
<td>1%</td>
<td>&gt;99.4%</td>
<td>&lt;0.6%</td>
<td>101 ± 2</td>
<td>3.6</td>
</tr>
<tr>
<td>CoOxd</td>
<td>&gt;99%</td>
<td>0.3 ± 0.7%</td>
<td>&lt;3%</td>
<td>95.9 ± 0.9%</td>
<td>4.1 ± 0.9%</td>
<td>63 ± 1</td>
<td>3.7</td>
</tr>
<tr>
<td>CoDyP</td>
<td>74 ± 1%</td>
<td>1.1 ± 0.4%</td>
<td>25 ± 1%</td>
<td>99.4 ± 0.4%</td>
<td>0.6 ± 0.4%</td>
<td>49 ± 0.2</td>
<td>6.5</td>
</tr>
</tbody>
</table>

*Measured by UPLC-MS.
†Measured by ICP-MS.
‡Concentration of “Co” (measured by ICP-MS) versus “protein” (measured by BCA assay).
§Calculated as milligram of holoprotein per liter culture. Yields were reproducible within a factor of 2.
CYP119 was rapid, occurring in less than 2 min. Interestingly, the visible spectrum and facile reduction behavior of Co(III)CYP119 were distinct from those of CoP450cam, which had been prepared through reconstitution of purified apoprotein with CoPPIX (49). In the absence of substrate, reduction of Co(III)P450cam required >60 min, but in the presence of substrate, complete reduction occurred within 10 min. Although CoPPIX-substituted Oxd has not been prepared previously, the native coordination is analogous to that of Mb, with a single His ligand bound to the heme iron. The absorption maxima of Co(III)Oxd are comparable to that of Co(III)Mb and Co(III)Mb*, suggesting that the native histidine is coordinated to cobalt. Reduction of Co(III)Oxd was slow and incomplete. A peak attributable to Co(III)Oxd remained at 429 nm after incubation with 4 mM sodium dithionite for 90 min at 35 °C. The absorption maxima for Co(III)Oxd were comparable to those of Co(II)Mb and Co(II)Mb*. The visible spectra of CoDyP are more complex due to the presence of significant amounts of H2PPIX-bound DyP. The Co(III)DyP sample exhibited absorption maxima at 618 nm and 664 nm, which are similar to Q bands previously observed for H2PPIX-bound TfDyP at 622 nm and 663 nm (76). However, the loss in intensity of these bands upon reduction suggests they might arise from charge transfer rather than from H2PPIX. Absorption maxima attributable to Co(III)DyP at 422 nm, 534 nm, and 567 nm are at comparable positions to those of Co(III)HRP, again suggesting retention of the native histidine ligand. Co(III)DyP reduced sluggishly to Co(II)DyP, at a rate similar to the reduction of CoMb* and requiring 45 min to reach completion. The resulting absorption maxima attributable to Co(II)DyP are similar to those of Co(II)Mb and Co(II)HRP.

We confirmed that native coordination to Co(II)PPIX was maintained in these proteins using EPR spectroscopy. Co(II)CYP119, Co(II)Oxd, and Co(II)DyP exhibit EPR spectra consistent with low-spin, five-coordinate, Co(II)porphyrin complexes, as indicated by the axial symmetry and eight hyperfine features due to coupling to the S = 7/2 59Co nucleus (Fig. 6). Spin Hamiltonian parameters derived from simulations of these spectra are provided in Table 3. The EPR spectrum of Co(II)CYP119 showed an overall spectral shape that was comparable to that observed in Co(II)P450cam. This spectrum is notably devoid of the triplet superhyperfine features that are associated with nitrogen coordination to cobalt and is therefore indicative of coordination by an I = 0 ligand such as sulfur. Notable are two distinct five-coordinate Co(II) signals: a major signal consistent with a stronger donor as the fifth ligand and a minor signal consistent with a weaker donor as the fifth ligand. These signals are analogous to those observed by Wagner et al. (49) in Co(II)P450cam, who attributed the major signal with a strong donor to axial thiolate ligation and the minor signal with a weak donor to thiol or water ligation. In contrast, CoCat and CoOxd exhibited triplet superhyperfine structure, conclusively demonstrating coordination by a nitrogen-bearing donor ligand. The presence of nitrogen superhyperfine coupling is consistent with the expected axial coordination by the native histidine residues in CoOxd and CoDyP. These EPR data are consistent with retention of the native coordination environment for Co(II)PPIX in the contexts of CoCYP119, CoOxd, and CoDyP.

In contrast to thiolate-coordinated CoCYP119 and histidine-coordinated CoOxd and CoDyP, expression of tyrosine-coordinated catalase failed to produce natively coordinated cofactor. Using the same expression conditions as previously, we successfully produced a CoPPIX-substituted catalase 3 from Neurospora crassa with reasonable yields and 54% of the protein loaded with cofactor (SI Appendix, Table S4). Of that cofactor, 99% was CoPPIX. However, spectroscopic characterization of the expressed protein revealed that the cobalt ion was not bound to the native tyrosinate ligand. Visible spectra of CoCat are shown in SI Appendix, Fig. S9, and peak maxima are provided in Table 2. No spectra of tyrosinate-ligated CoPPIX are available for comparison, but the peak maxima of Co(II)Cat (Table 2) are comparable to those of CoMb, suggestive of histidine coordination to CoPPIX. The EPR spectrum of Co(II)Cat is highly informative and exhibits two distinct signals (SI Appendix, Fig. S9). The first signal is axially symmetric, consistent with a five-coordinate cobalt complex, and exhibits poorly resolved triplet superhyperfine structure, indicative of a nitrogen-containing ligand bound to Co(II). The second signal, the significant feature
centered around $g = 2$ that exhibits cobalt hyperfine coupling, is suggestive of a superoxide complex of Co(III), which may have formed upon reaction with oxygen during sample transfer. The shapes and $g$ values for this latter signal are reminiscent of those observed for CoMBO$_2$ (4). Importantly, neither of these signals are what would be expected for tyrosinate coordination. Interrogation of the crystal structure of \textit{N. crassa} catalase 3 (Protein Data Bank [PDB] 5WHQ) reveals that His-102 resides near the heme iron on the distal side opposite the native tyrosine ligand. It is therefore plausible that in the CoPPIX-substituted protein, this histidine coordinates to cobalt in place of the native tyrosine. Failure to bind the native phenolate ligand is unsurprising given the poor oxophilicity of cobalt. Thus, ligation of cobalt-substituted hemoproteins produced via this method is expected to follow the chemical propensity of the cobalt center.

**Conclusions**

We serendipitously discovered that common \textit{E. coli} BL21(DE3) is able to produce a new-to-nature CoPPIX cofactor without genetic engineering, evolutionary adaptation, or auxiliary plasmids and can efficiently incorporate this cofactor into hemoproteins. The toxicity of free cobalt, which is present in high concentrations in the growth medium, is partly offset by overexpression of the hemoprotein. We surmise that this enhanced cobalt tolerance is due to sequestration of toxic cobalt within the hemoprotein in the form of CoPPIX. To our surprise, cells rendered tolerant to cobalt by serial passaging were less efficient at incorporating CoPPIX, and, in fact, selection for cobalt tolerance appeared to enhance incorporation of the native FePPIX. When plasmid-bearing \textit{E. coli} BL21(DE3) cells are grown in augmented minimal media supplemented with 500 μM CoCl$_2$, CoPPIX is readily biosynthesized and incorporated into a variety of heterologously expressed hemoproteins with an efficiency that is comparable to their native FePPIX cofactor. Our analysis demonstrates that CoPPIX is usually incorporated into the native binding site in the expressed protein, with only trace amounts of FePPIX observed.

While the method here is only demonstrated with cobalt, future studies may explore the introduction of alternative metals. Regardless, the utility of cobalt substituted hemoproteins is broad. CoPPIX is an excellent spectroscopic analog of heme that has been used extensively in EPR and resonance Raman studies to characterize hemoproteins. Notably, the study of synthetic cobalt porphyrin complexes suggests there is a wealth of chemistry that remains to be explored and expanded upon in a biocatalytic setting (30–32). Cobalt itself is attractive as it is abundant and inexpensive compared to many transition metals. This route to CoPPIX synthesis and biocorporation leverages a common laboratory strain and requires no auxiliary plasmids. Protein expression does not require specialized equipment, anaerobic culture conditions, or supplementation of a presynthesized cofactor. Consequently, this method can be applied using common resources by any research group equipped for protein expression in \textit{E. coli}. Furthermore, the straightforward nature of this approach lends itself to directed evolution, which we speculate may unlock as yet untapped modes of biocatalytic CoPPIX reactivity.

**Materials and Methods**

Detailed experimental methods can be found in SI Appendix.

**Protein Expression and Purification.** Cobalt-substituted hemoproteins were expressed in \textit{E. coli} BL21(DE3) bearing a pET22b plasmid encoding the protein of interest. A modified M9 media (M9*) was used for the overnight starter culture, which was used to inoculate 1 L of M9* expression culture. CoCl$_2$ (500 μM) and G6A (250 μM) were added to the expression culture once OD$_{600}$ reached 0.2 to 0.3. Once OD$_{600}$ reached 0.6 to 0.7, expression was induced with the addition of IPTG (250 μM). After overnight incubation, cells were harvested by centrifugation and stored at −20 °C until purification. Following lysis by sonication, the cell lysate supernatant was purified by nickel affinity chromatography. Additional details can be found in SI Appendix.

**Quantitation of Cobalt Tolerance with IC$_{50}$ Assays.** To measure the ability of various \textit{E. coli} strains and expression systems to grow in cobalt-supplemented M9* media, a 96-well plate-based IC$_{50}$ assay was used. Briefly, single colonies were used to inoculate a 5 mL M9* overnight starter culture. These starter cultures were in turn used to inoculate (2.5% inoculum) 200 μL of M9* media in each well of a 96-well plate, containing variable amounts of CoCl$_2$. After 9 h, the OD$_{600}$ of each well was measured using a plate reader. GraphPad PRISM software was used to calculate the cobalt IC$_{50}$ value for each \textit{E. coli} strain or expression system. Additional experimental details and data analysis methods can be found in SI Appendix.

**Electronic Absorption Spectroscopy.** Electronic absorption spectra were taken of purified protein samples in argon-sparged 50 mM potassium phosphate buffer. Spectra of reduced protein samples were taken after reduction with

---

**Table 2. Comparison of electronic absorption maxima of CoPPIX-substituted hemoproteins produced in vivo with analogous proteins produced through in vitro reconstitution with CoPPIX**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Ligands</th>
<th>Soret ($\lambda$)</th>
<th>$\beta$</th>
<th>$\alpha$</th>
<th>Ligand</th>
<th>Soret ($\lambda$)</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb</td>
<td>His/H$_2$O</td>
<td>424</td>
<td>535</td>
<td>572</td>
<td>His</td>
<td>406</td>
<td>558</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>MbH64V V68A (Mb*)</td>
<td>His/H$_2$O</td>
<td>428</td>
<td>540</td>
<td>572</td>
<td>His</td>
<td>406</td>
<td>555</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>MbH64V V68A (Mb*)</td>
<td>His/H$_2$O</td>
<td>427</td>
<td>536</td>
<td>577</td>
<td>His</td>
<td>406</td>
<td>557</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>P450$_{cam}$</td>
<td>Cys/H$_2$O</td>
<td>422</td>
<td>538</td>
<td>570</td>
<td>Cys</td>
<td>404</td>
<td>556</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>CYP119</td>
<td>Cys/H$_2$O</td>
<td>420</td>
<td>539</td>
<td>566</td>
<td>Cys</td>
<td>410</td>
<td>558</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>Oxd</td>
<td>His/H$_2$O</td>
<td>429</td>
<td>542</td>
<td>574</td>
<td>His</td>
<td>402</td>
<td>559</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>HRP(B,C)</td>
<td>His/H$_2$O</td>
<td>421</td>
<td>533</td>
<td>565</td>
<td>His</td>
<td>401</td>
<td>553</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>DyP</td>
<td>His/H$_2$O</td>
<td>422</td>
<td>534</td>
<td>567</td>
<td>His</td>
<td>403</td>
<td>564</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>Cat</td>
<td>N?</td>
<td>427</td>
<td>537</td>
<td>570</td>
<td>N</td>
<td>403</td>
<td>561</td>
<td>This work</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3. Comparison of EPR parameters for Co(II)PPIX-substituted hemoproteins produced using in vivo expression, with those of analogous proteins produced using in vitro reconstitution with CoPPIX**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Effective g value</th>
<th>$A_C$</th>
<th>$A_A$</th>
<th>$\beta$</th>
<th>$\alpha$</th>
<th>$k$</th>
<th>$\beta$</th>
<th>$k$</th>
<th>$\alpha$</th>
<th>$k$</th>
<th>$\beta$</th>
<th>$k$</th>
<th>$\alpha$</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mb</td>
<td>3.23 2.03 6</td>
<td>79</td>
<td>NR</td>
<td>17</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mb H64V V68A (Mb*)</td>
<td>3.22 2.04 2.4</td>
<td>77.6</td>
<td>2.2</td>
<td>16.9</td>
<td>This work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P450$_{cam}$</td>
<td>3.22 2.03 2.7</td>
<td>70.2</td>
<td>2.2</td>
<td>16.9</td>
<td>This work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP119</td>
<td>3.22 2.03 2.7</td>
<td>70.2</td>
<td>2.2</td>
<td>16.9</td>
<td>This work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxd</td>
<td>3.21 2.02 7.2</td>
<td>83.4</td>
<td>2.3</td>
<td>16.7</td>
<td>This work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRP (A)</td>
<td>3.24 2.03 11.5</td>
<td>73.9</td>
<td>NR</td>
<td>17.3</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DyP</td>
<td>3.23 2.04 10.4</td>
<td>73.1</td>
<td>6.6</td>
<td>16.7</td>
<td>This work</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NR, not reported.
sodium dithionite (3 to 4 mM). Additional information and instrument parameters can be found in SI Appendix.

**EPR Spectroscopy.** EPR spectra of reduced, purified protein were taken in 50 mM potassium phosphate buffer. Sodium dithionite was used as the reductant, and complete reduction of each sample was confirmed with electronic absorption spectroscopy. Samples were frozen in liquid nitrogen following preparation in an anaerobic chamber. Additional information and instrument parameters can be found in SI Appendix.

**Quantitation of Porphyrin Content.** The metal content for purified protein samples was measured using ICP-MS and comparison to a standard curve.

---

**Fig. 5.** Characteristics of hemoproteins chosen for CoPPix-substituted hemoprotein production. Each protein is listed with name, abbreviation, native organism, and biological function. Structures of CYP119 (PDB: 1F4U), Oxd (PDB: 3A15), and DyP (PDB: 5VJ0) are shown as cartoons highlighting secondary structure. Structures of heme b in each protein highlighting axial ligation are shown. In all structures, iron is shown as an orange sphere, and protoporphyrin IX and axial ligands are shown as sticks.

---

**Fig. 6.** Spectroscopic characterization of other CoPPix-substituted hemoproteins produced by overexpression in CoCl₂-supplemented media. (A–C) Electronic absorption spectra of CoCYP119, CoOxd, and CoDyP at 35 °C in Ar-sparged 50 mM potassium phosphate buffer, pH 7. Co(III) spectra are shown in black. Co(II) spectra are shown in color. Reduction was initiated by addition of Na₂S₂O₄ to a final concentration of 3 to 4 mM. Spectra tracking the progress of reduction are shown as gray dotted lines. No intermediate spectra are shown for CoCYP119 as complete reduction occurred prior to the first spectral scan after addition of reductant. (D–F) EPR spectra of Co(II) hemoproteins. Darker colors are experimental spectra. Lighter colors are simulated spectra. Asterisks denote a consistent signal derived from sample cavity. Parameters from spectral simulations are provided in Table 2. Acquisition parameters are provided in SI Appendix, General Experimental Methods.
The relative porphyrin content of cell pellets and purified proteins was assessed using UPLC-MS analysis of acidified pyridine extracts and comparison to a standard curve. The protein content of purified protein samples was assessed using a BCA assay. Additional details for each of these methods can be found in SI Appendix. Data Availability. All study data are included in the article and/or SI Appendix.

ACKNOWLEDGMENTS. We acknowledge the invaluable support, assistance, and advice from our colleagues in the Buller and Burstein laboratories in particular, we would like to thank Dr. Matt Dent and Prof. Thomas Brunold for their expertise. This work is funded by a PEARL fellowship funded by the Qatar National Research Fund (QNRF) and the National Science Foundation (NSF) grant to PMD. We thank Dr. Pam Doolittle for her assistance with ICP-MS analysis. We also acknowledge Shimadzu Instruments, Columbus, MD. The ICPSM-2030 used in this study was acquired utilizing a grant provided as part of Shimadzu’s Academic Grant Program. We acknowledge their support. Support for this research was provided by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation. B.R.W. was supported by NSF Graduate Research Fellowship Program under Award Number DGE-1747503.


62. J. Sudhamsu et al., Co-expression of ferrochelatase allows for complete heme incorporation into recombinant proteins produced in E. coli. Protein Expr. Purif. 73, 78–82 (2010).


68. V. Haskamp et al., The radical SAM protein HemW is a heme chaperone. J. Biol. Chem. 293, 2558–2572 (2018).

69. E. A. Sweeny et al., Glyceraldehyde-3-phosphate dehydrogenase is a chaperone that allocates labile heme in cells. J. Biol. Chem. 293, 14557–14568 (2018).


