# Allosteric cooperation in a de novo-designed two-domain protein

Fabio Pirro[a,1], Nathan Schmidt[b,1], James Lincoff[b], Zachary X. Widel[c], Nicholas F. Polizzi[b], Lijun Liu[d,e,2], Michael J. Therien[c], Michael Grabe[b], Marco Chino[a], Angela Lombardi[a,3], and William F. DeGrado[b,3]

[a]Department of Chemical Sciences, University of Napoli Federico II, 80126 Napoli, Italy; [b]Department of Pharmaceutical Chemistry and the Cardiovascular Research Institute, University of California, San Francisco, CA 94158-9001; [c]Department of Chemistry, Duke University, Durham, NC 27708-0346; [d]State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, 518055 Shenzhen, China; and [e]DLX Scientific, Lawrence, KS 66049

Contributed by William F. DeGrado, October 29, 2020 (sent for review August 20, 2020; reviewed by Giovanna Ghirlanda and Shun Hirota)

**We describe the de novo design of an allosterically regulated protein, which comprises two tightly coupled domains. One domain is based on the DF (Due Ferri in Italian or two-iron in English) family of de novo proteins, which have a diiron cofactor that catalyzes a phenol oxidase reaction, while the second domain is based on PS1 (Porphyrin-binding Sequence), which binds a synthetic Zn-porphyrin (ZnP). The binding of ZnP to the original PS1 protein induces changes in structure and dynamics, which we expected to influence the catalytic rate of a fused DF domain when appropriately coupled. Both DF and PS1 are four-helix bundles, but they have distinct bundle architectures. To achieve tight coupling between the domains, they were connected by four helical linkers using a computational method to discover the most designable connections capable of spanning the two architectures. The resulting protein, DFP1 (Due Ferri Porphyrin), bound the two cofactors in the expected manner. The crystal structure of fully reconstituted DFP1 was also in excellent agreement with the design, and it showed the ZnP cofactor bound over 12 Å from the dimetal center. Next, a substrate-binding cleft leading to the diiron center was introduced into DFP1. The resulting protein acts as an allosterically modulated phenol oxidase. Its Michaelis–Menten parameters were strongly affected by the binding of ZnP, resulting in a fourfold tighter $K_m$ and a 7-fold decrease in $k_{cat}$. These studies establish the feasibility of designing allosterically regulated catalytic proteins, entirely from scratch.**

de novo design | diiron protein | porphyrin-binding protein | allostery | protein evolution

The emergence of life and the evolution of the three superkingdoms required the recombination of preexisting protein domains to perform ever-increasingly complex functions (1). The majority (~90%) of multidomain proteins are made up of end-to-end linked domains; in the remaining ~10% of cases, a domain insertion occurs, creating a continuous and a discontinuous domain (2, 3). In enzymes and small molecule binding proteins, the bilobed architecture facilitates the formation of active sites between individual domains; redox-active proteins often combine multiple domains to orient multiple cofactors for productive electron transfer; and allosterically regulated proteins combine binding domains with catalytic or signal-transducing domains (4). Therefore, the addition of a domain to an existing protein expands, alters, or modulates its functionality (5). Protein engineers have been inspired by this modularity to generate artificial multidomain proteins with improved properties or to create nanostuctured and sensor devices (6–10). Toward this end, different methodologies have been developed to fuse the different domains by: 1) introducing designed or naturally occurring peptide linkers (11); 2) superimposing and fusing one or two turns of terminal alpha helices of connecting helical proteins (12, 13); and 3) computationally designing new structural elements to interface the different domains in a fragment based approach (14). In each case, the domain architectures of the artificial multidomain proteins fell into either end-to-end or domain insertion topology. However, the de novo design and structure determination of allosterically regulated multidomain proteins (in which both domains are designed from scratch) have not been reported.

Here, we describe the design of a protein that combines domains capable of binding ZnP, the Zn[5,10,15,20-tetrakis(trifluoromethyl) porphinato], and diiron cofactors into a single tightly coupled framework. While multicofactor proteins have been widely used to explore redox coupling (15–22), high-resolution structures of multicofactor proteins have not been described in the literature, limiting what can be learned and achieved in such systems. The present work differs in two fundamental manners from earlier work on the design of multicofactor proteins. First, the goal of the present study was to examine how the binding of redox-inert ZnP cofactor allosterically modulates the catalytic activity of a second diiron-binding domain. Second, the structure of the designed multicofactor protein was determined by X-ray crystallography. It is also noteworthy that the computational methods, adopted in this work, could be readily extended to the design of electronically coupled systems for light-triggered electron energy storage and utilization, particularly given the ability to design proteins that incorporate metal ion clusters (15, 21, 23–31).

The diiron-binding component of our two-domain protein is based on the DF family of de novo proteins (32, 33), which have been optimized to catalyze various two and four-electron reactions, including ferroxidase, oxidase, and monooxygenase activities (34–40). The second domain is based on PS1, which binds the

## Significance

**A major mechanism of evolution involves fusing genes that encode single-domain proteins to create multidomain structures that achieve new functions. Here, we develop methods to design multidomain proteins entirely from scratch and achieve the premier de novo design of an allosterically regulated phenol oxidase that responds to the binding of a synthetic porphyrin.**

Author contributions: F.P., N.S., J.L., Z.X.W., N.F.P., L.L., M.J.T., M.G., M.C., A.L., and W.F.D. designed research; F.P., N.S., J.L., Z.X.W., L.L., and M.C. performed research; F.P., J.L., N.F.P., L.L., M.J.T, M.C., A.L., and W.F.D. analyzed data; and F.P., N.S., J.L., Z.X.W., N.F.P., L.L., M.J.T., M.G., M.C., A.L., and W.F.D. wrote the paper.

Reviewers: G.G., Arizona State University; and S.H., Nara Institute of Science and Technology.

The authors declare no competing interest.



[1]F.P. and N.S. contributed equally to this work.

[2]Present address: Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, KS 66047.

[3]To whom correspondence may be addressed. Email: alombard@unina.it or william.degrado@ucsf.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2017062117/-/DCSupplemental.

First published December 14, 2020.

synthetic ZnP (41). PS1 has a well-structured hydrophobic core, which positions the porphyrin-binding domain for productive interaction with this cofactor. The structure of PS1 has been solved by NMR in both the apo-bound and ZnP-bound state. The structures are nearly identical in the hydrophobic core, but the apo-protein is more open and flexible near the opposite end of the bundle. This structural transition allows the protein to bind the cofactor in an alligator-like chomping motion. We reasoned that such transition might be used to regulate the properties of the DF diiron site in the neighboring domain. Both DF and PS1 are four-helix bundles, so we envisioned a coaxial arrangement to facilitate interdomain communication, as in bacterial signaling proteins (42). However, DF and PS1 have distinct bundle architectures with respect to their interhelical packing, helical offsets, and helical registers, which together presented challenges for structural design. To address these challenges, we extended previous fragment-based approaches (43–48) and designed artificial multidomain proteins with allosterically communicating sites.

## Results and Discussion

**Design Strategy.** The DF and PS1 scaffolds differ in the interhelical separation and registry of their four-helix bundle domains (*SI Appendix*, Fig. S1). Therefore, it was necessary to design linkers that connect the two bundles in an energetically reasonable manner. To accomplish this, the helical bundles were connected using a structure-based design strategy involving two interrelated steps: first, the optimal relative translation and orientation between the bundles, namely the bundle placement, were determined; second, a structural solution interconnecting each of the disjointed helices between the domains was identified.

In the most general case, the bundle placement involves searching a six-dimensional space, comprising three translational

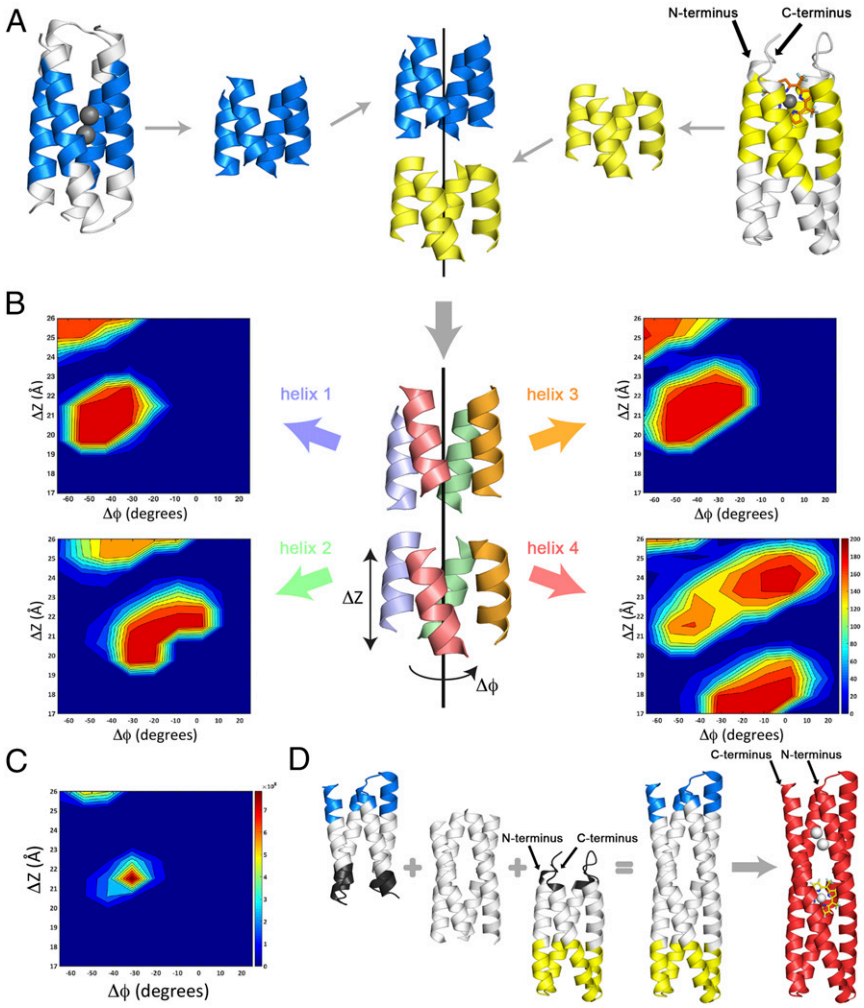



**Fig. 1.** Overview of the design strategy, the structural designability landscapes, and construction of the helical bundle design. (*A*) The helical bundles are first aligned along a common axis. In both bundles, the helical segments surrounding the dimetal and porphyrin ligands are identified (colored in blue and yellow, respectively) and used to determine their bundle axis. The coaxial bundle alignment reduces the relative positioning of the bundles down to two degrees of freedom: the relative translation along the bundle axis, $\Delta Z$, and relative rotation about the bundle axis, $\Delta\Phi$. (*B*) The structural designability of the four pairs of helices between bundles is determined. The bundle (*Middle*) is decomposed into four searches for each disjointed helical pair, helices 1, 2, 3, and 4 (same color), which are considered separately (the four panel corners). In each search one helix is fixed (*Top*) and the other is translated, $\Delta Z$, and rotated, $\Delta\Phi$, about the bundle axis (depicted on bottom). Structural matches are then tallied for each coordinate position, ($\Delta Z$,$\Delta\Phi$) and the resulting 2D plot describes the designability "landscape" of the helical pair. The four helix designability landscapes show designable "hotspots" with ~200 structural matches surrounded by vast nondesignable regions with 0 structural matches (scale included in helix 4 landscape). (*C*) The bundle designability landscape obtained by multiplying the number of structural matches from the four designability landscapes for each helical pair. Maxima in the bundle designability landscape indicate optimal relative positioning of the two bundles. (*D*) Final construction of the helical bundle. The bottom/top parts of the metal/porphyrin-binding proteins (black) are removed, while the top/bottom portions (blue/yellow) are reconnected to the joined, designed helices from the two bundles. The resulting DFP1 design, with metal and porphyrin ligands, is shown at the far right.

and three rotational degrees of freedom. In many helical bundles (49), including the ones considered here (33, 41), the helices supercoil about a central or bundle axis. By requiring the bundles to share a common bundle axis (Fig. 1*A*), the search space is reduced to two dimensions: the translational displacement along the bundle axis, ΔZ, and the rotational displacement about the bundle axis, ΔΦ. This approach requires locating each bundle axis, which is ultimately set by the helical chains comprising each respective bundle. Once determined (*SI Appendix*), the colinear alignment of the two bundle axes and the relative bundle positioning (ΔZ, ΔΦ) are achieved by applying the appropriate rotation and translation matrices to the atomic coordinates of each structure.

With this parameterization, the search for the optimal bundle placement became a search through (ΔZ, ΔΦ) space, which we systematically vary to discover poses in which all four gaps in the helices can be simultaneously spanned with helical linkers. We examined strategies in which the linkers were selected from the database individually or in pairs. We first examined the four backbone connections between bundles separately. The $i^{th}$ bundle pose, defined by coordinates ($\Delta Z_i$, $\Delta\Phi_i$), was decomposed into four poses corresponding to each pair of disjointed helices from the two bundles (Fig. 1*B*). Poses were scored using a structural designability criterion implemented through the structural similarity search program MASTER (48). In MASTER, an input structural element "query" is compared to a collection of three-dimensional protein structures "library" using a similarity metric of the backbone rmsd (<1.0 Å). Library structural fragments that match the query are enumerated and the designability of a query is quantified by the total number of structural matches found in the library. Operationally, for coordinates $\Delta Z_i$, $\Delta\Phi_i$, the four poses from each pair of disjointed helices became four distinct queries that were scored by their designabilities, $H_n(\Delta Z_i, \Delta\Phi_i)$, with $n = 1$, 2, 3, 4. The search through (ΔZ, ΔΦ) space was readily visualized by a two-dimensional (2D) contour plot (Fig. 1*B*), with designable regions indicated as peaks surrounded by less designable valleys. To recombine these four distinct searches into the bundle designability landscape B(ΔZ, ΔΦ), we took the product of the four backbone designabilities: $B(\Delta Z, \Delta\Phi) = H_1(\Delta Z, \Delta\Phi)\times H_2(\Delta Z, \Delta\Phi)\times H_3(\Delta Z, \Delta\Phi)\times H_4(\Delta Z, \Delta\Phi)$ for all ΔZ, ΔΦ (Fig. 1*C*). The maxima in B(ΔZ, ΔΦ) determined the optimal bundle placement and, thereby, position the four disjointed helical pairs, each of which had to be connected by a polypeptide chain.

**Structural Model.** The bundle designability landscape showed a single maximum "hotspot" at (ΔZ, ΔΦ) = (21.5 Å, −33.8°) (Fig. 1*C*) corresponding to bundle placement in which each pair of helices was offset in both translation and rotation to accommodate the gentle left-handed supertwist of the bundles (*SI Appendix*, Fig. S2*A*). Inspection of the designability landscapes of the four backbone connections revealed the positions of their hotspots had some degree of similarity (Fig. 1*B* and *SI Appendix*, Fig. S2*B*), yet they also differed due to underlying differences between bundle geometries (*SI Appendix*, Figs. S2–S5). Analogous designability searches considering each adjacent pair of helices in the bundle (i.e., helices 1 & 2, 2 & 3, 3 & 4, and 4 & 1), showed the same trends (*SI Appendix*, Fig. S3). Structural matches occurred when the translational and rotational offset between helical fragments were compatible with polypeptide linkers that adopted designable helix geometries. Moreover, the flexibility in those geometries allowed simultaneous structural matches to be realized in the four helices in the bundle.

In particular, the coordinates of the maxima in the bundle designability plot (Fig. 1*C*) identified the most favorable set of linker geometries (indicated by black dots in *SI Appendix*, Fig. S2*B*). The top structural matches to the disjointed helices (*SI Appendix*, Fig. S2*C*) all displayed helical geometry, with distortion from ideality in two cases (*SI Appendix*, Fig. S5). Helices 1–3 were best connected with a two-residue helical linker, whereas the structural matches for helix four consisted of five-residue and six-residue (*SI Appendix*, Fig. S2*C*, white, and *SI Appendix*, Fig. S4) linkers. To construct the final backbone structure, the four helical backbone fragments were connected into a single chain by incorporating the N and C termini plus the loop from the DF structure (Fig. 1*D*, colored blue), and the loops and helical regions comprising the folded core from the porphyrin-binding structure (Fig. 1*D*, colored yellow). Sequence design was restricted to the helical segments where the distinct bundles were connected, at residue positions that do not have side chains within the first and second shell of the dimetal-binding and porphyrin-binding sites. Backrub within Rosetta was used to sample small structural changes around the connections in conjunction with alternating loops of fixed-backbone sequence design and backbone/sidechain minimization (*SI Appendix*). The results from Rosetta were compared with sequence information from the MASTER searches (*SI Appendix*, Fig. S5), and these data sources along with visual inspection of the model were used to finalize the primary structure.

**Structural Characterization of DFP1.** DFP1 was cloned and expressed as described in *Materials and Methods*. For spectroscopic and structural characterization, we reconstituted the protein with ZnP in the porphyrin-binding site. $Zn^{2+}$ was introduced into the dimetal site as a redox-inactive stable mimic of ferrous ions, as in previous studies of DF proteins. Spectral titration with ZnP demonstrated that the protein bound tightly to the cofactor. Time-resolved transient absorption spectroscopic data acquired for ZnP-DFP1 (*SI Appendix*, Fig. S6*A*) evinces spectral features and excited-state dynamics characteristic of the benchmark ZnP-PS1 holoprotein (41). Identical experiments carried out with di-$Zn^{2+}$-ZnP-DFP1 (*SI Appendix*, Fig. S6*B*) indicate that $Zn^{2+}$ occupancy of protein dimetal binding site does not perturb the excited-state relaxation dynamics of the ZnP chromophore or its characteristic electronically excited singlet ($S_1$)- and triplet ($T_1$)-state absorptions. Because $S_1\rightarrow T_1$ intersystem crossing rate constants of electronically excited porphyrins are known to be sensitive to both macrocycle structure and the local environment (50), these experiments demonstrate that the ZnP-binding site in PS1 is faithfully reproduced in DFP1.

The structure of the holoprotein was solved by molecular replacement to 3.5-Å resolution (Fig. 2*A*). At this resolution, the metal ions are very well resolved, and the density is sufficiently clear, as well as the OMIT maps, to allow placement of the porphyrin macrocycle and the zinc ions (*SI Appendix*, Figs. S7 and S8). The closest distance between a porphyrin carbon and a $Zn^{2+}$ ion of the dimetal cofactor is 12 Å. Furthermore, we investigated whether the individual domains of DFP1 preserved the structural features of the starting single-domain proteins. Indeed, we observed an excellent agreement considering individually the two domains: 0.74 Å rmsd for superposition of the ZnP domain versus PS1 and 0.52 Å for the dimetal binding domain versus DF1 (*SI Appendix*, Fig. S9). Moreover, the superposition of the 120 residues comprising the cofactor binding sites gives an excellent fit to the design model with an rmsd = 0.8 Å (Fig. 2 *B* and *C*). At this resolution, it was not possible to detect deviations from ideal hydrogen-bonded geometries in the helices, but each connecting helix showed continuous density with B-factors within the range seen for the main chain atoms in other portions of the protein.

**Design and Solution Characterization of Catalytically Active *DFP3*.** DFP1 was designed for maximal thermodynamic stability, and its interior is well-packed with apolar sidechains throughout the bundle. The resulting tight and uniform packing provided high stability, but did not leave room for organic substrates to access the dimetal-binding site. Therefore, the four interacting Leu and Ala residues located just above the dimetal site were substituted to Gly residues (Fig. 3 *A* and *B*, respectively), resulting in a deeply invaginated substrate access cavity. The resulting four-site mutant with Gly substitutions at positions 10, 14, 71, and 74

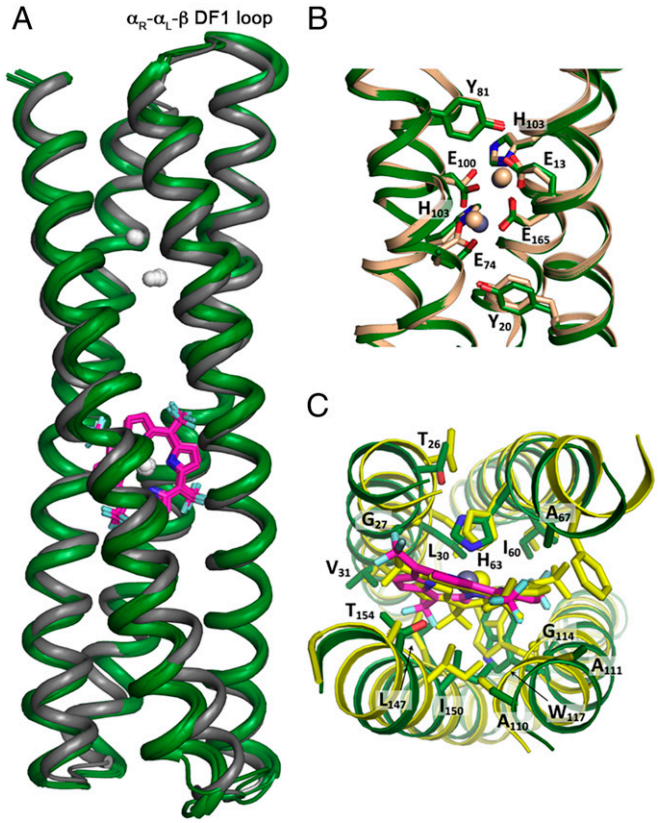



**Fig. 2.** The crystal structure of DFP1 displays Å-level agreement with the design model and previous dimetal- and porphyrin-binding helical bundles. (*A*) Structures from the four chains in the asymmetric unit (green) superimposed onto the design (gray). $Zn^{2+}$ ions in the dimetal site and in ZnP are shown as white spheres. The porphyrin macrocycle is shown as pink sticks. (*B*) Superimposition around the dimetal center of the crystal structures of di-$Zn^{2+}$-ZnP-DFP1 (in green, PDB ID code: 7JH6) with di- $Zn^{2+}$-DF1 (tan, PDB ID code: 1EC5). (*C*) Top-down view comparing the porphyrin binding sites of di-$Zn^{2+}$-ZnP-DFP1 and ZnP-PS1 (yellow, PDB ID code: 5TGY). Key first-shell and second-shell residue side chains are shown as sticks and numbered according to the DFP1 sequence.

(designated DFP2, *SI Appendix*, Fig. S10) was expressed and remained folded despite the presence of four helix-destabilizing mutations. Nevertheless, it was found to exist in a monomer-dimer equilibrium, similar to what had been observed for DFP1. Size exclusion chromatography (SEC) showed the presence of two peaks of approximately equal integrated intensity at roughly the positions expected for a monomer and dimer. Upon addition of a single equivalent of the porphyrin, both peaks sharpened and shifted to a longer retention time (Fig. 3*E*), indicative of a more compact conformation. Additionally, the intensity of the monomeric peak increased, although dimers still accounted for about 20% of the total.

We hypothesized that the dimeric peak might represent an elongated domain-swapped dimer of the desired protein, as previously observed in several cytochromes and myoglobins (51–54). The dimeric conformation might have similar binding properties, but it would complicate solution analysis. To address this problem, we introduced a helix-breaking sequence that could stabilize the intended $\alpha_R$-$\alpha_L$-β interhelical loop used in DFP1. Based on a previous survey of the Protein Data Bank (PDB), the Val-Lys-Leu at positions 88–90 of DFP1 was substituted to Thr-His-Asn in DFP3 (Fig. 3 *C* and *D*, respectively) (55). The resulting protein, DFP3, formed only a small amount of dimer in the apo-state and when bound in the ZnP-bound state. (Fig. 3*F* and *SI Appendix*, Fig. S10). Moreover, the retention times were identical and as expected from the crystal structure of DFP1 (*SI Appendix*, Table S2). This finding strongly suggested that the DFP3 was largely folded and preorganized in the absence of cofactor.

**Analysis of Binding of Porphyrin and Metal Ions to DFP3.** To evaluate the binding of ZnP to DFP3, we took advantage of the spectral shifts in the ZnP Soret band previously seen in PS1. The spectrum of ZnP solubilized in 1% wt/vol octyl-β-D-glucopyranoside micelles sharpened and shifted from 415 to 423 nm when the cofactor was titrated with increasing concentrations of DFP3 (Fig. 4*A*). Nonlinear least squares fitting of the resulting spectral titration revealed a 1:1 binding stoichiometry with an apparent dissociation constant of $K_D = 10 \pm 2$ nM. This value represents an upper limit of the true binding constant, as it reflects the process of transfer of the porphyrin derivative from a micelle to the protein interior rather than from water to protein. A similar $K_D = 13 \pm 3$ nM was observed in the presence of excess $Zn^{2+}$ to

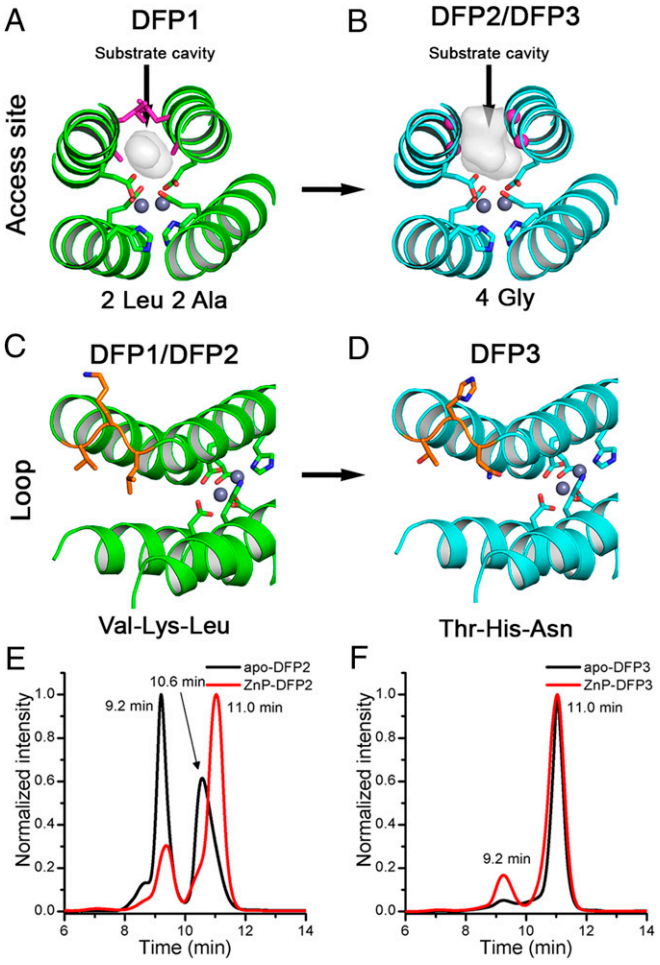



**Fig. 3.** Engineering functionality in DFP family. The substrate cavity at the dimetal site was broadened mutating the 2 Leu and 2 Ala residues in DFP1 (shown as magenta sticks in *A*) in 4 Gly residues in DFP2 and DFP3 (shown as magenta spheres in *B*). To improve the oligomeric behavior, we mutated the Val-Lys-Leu loop in DFP1 and DFP2 (shown as orange sticks in *C*) in Thr-His-Asn in DFP3 (shown as orange sticks in *D*). The backbone is shown as green and cyan cartoon before and after the mutations, respectively. The coordinating residues of the dimetal center are shown as sticks, and the metal ions in gray spheres. (*E*) SEC of apo-DFP2 (in black) and ZnP-DFP2 (in red) at pH 7 (Hepes 50 mM, NaCl 100 mM), followed at 280 nm. (*F*) SEC of apo-DFP3 (in black) and ZnP-DFP3 (in red) at pH 7 (Hepes 50 mM, NaCl 100 mM), was followed at 280 nm. Retention time peaks at 10.6 min and 11.0 min corresponded to hydrodynamic radii of 27 Å and 21 Å, respectively.

saturate the dimetal-binding site (*SI Appendix*, Fig. S11). Thus, DFP3 binds ZnP with exceptionally high affinity both in the apo and in the dimetal occupied states.

Although ZnP is achiral and, hence, not optically active, its ruffled conformation is rendered chiral when bound with an axial His ligand in the asymmetric binding site environment, as observed in DFP1 crystal structure. To confirm this feature in the related DFP3 protein, we examined the circular dichroism (CD) of its Soret electronic transition. The ZnP-DFP3 complex displayed a very intense negative Cotton effect at ~415 nm (Fig. 4*B*), which was not observed for ZnP dissolved in organic solution, either in the presence or absence of a histidine derivative (*SI Appendix*, Figs. S12 and S13). These findings confirm that ZnP is bound in a relatively rigid asymmetric environment both in DFP1 and DFP3 proteins.

While the crystal structure of holo-DFP1 demonstrated the binding of two $Zn^{2+}$ in the expected geometry, we also wished to probe the ability of apo-DFP3 to bind the first-row transition metal $Co^{2+}$ in aqueous solution. The position and intensity of the d-d bands of $Co^{2+}$ ions are very sensitive to the stoichiometry and geometry of the ions' ligand environment, providing a convenient probe of ligation in solution (56). The observed molar extinction coefficients ($\varepsilon_{524nm}$ = 146 $M^{-1}{\cdot}cm^{-1}$; $\varepsilon_{547nm}$ = 151 $M^{-1}{\cdot}cm^{-1}$; $\varepsilon_{600nm}$ = 82 $M^{-1}{\cdot}cm^{-1}$) of the di-$Co^{2+}$-DFP3 complex are indicative of a pentacoordinate geometry, as seen in the earlier observed di-$Co^{2+}$-DF complexes (Fig. 4*C*) (33, 34, 38, 57). Unfortunately, the presence of the ZnP Q-bands in region between 550 nm and 620 nm did not allow the detection of the cobalt binding in presence of the ZnP (absorption coefficient of Q bands and cobalt d-d transition, ~12,000 $M^{-1}{\cdot}cm^{-1}$ and 150 $M^{-1}{\cdot}cm^{-1}$, respectively).

Finally, we examined the influence of cofactor binding on the thermodynamic stability of the protein to heat and guanidine hydrochloride (Gdn-HCl)-induced unfolding. The apo-form of DFP3 shows a far ultraviolet (UV) CD spectrum typical of an α-helical protein (Fig. 4*D*), and its spectrum is exceptionally stable to thermal denaturation, showing only the beginning of an unfolding transition near 90 °C. The di-$Zn^{2+}$, ZnP, and doubly loaded di-$Zn^{2+}$ -ZnP proteins showed nearly identical far UV spectra with high stability to thermal unfolding up to 95 °C (Fig. 4*D* and *SI Appendix*, Fig. S14*A*). To resolve the differences in stabilities of the various forms of the protein, thermal unfolding was repeated in the presence of denaturant, 4 M Gdn-HCl (*SI Appendix*, Fig. S14*B*). The observed midpoints of the unfolding curves were as follows: di-$Zn^{2+}$-DFP3 = 71.0 ± 0.2 °C, ZnP-DFP3 = 83.2 ± 0.7 °C, and di-$Zn^{2+}$-ZnP-DFP3= 87.3 ± 0.4 °C. The exceptional stability is particularly important for future studies that might require destabilizing substitutions to be introduced to facilitate function.

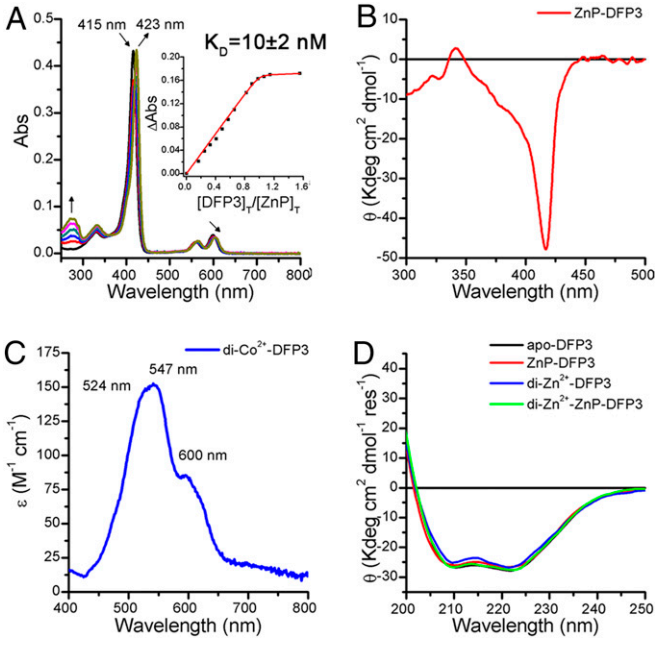



**Fig. 4.** DFP3 binds the single metal cofactors as the parental proteins. (*A*) UV-vis spectral changes of a 2.2 μM ZnP solution upon addition of apo-DFP3 at pH 7 (Hepes 50 mM, NaCl 100 mM octyl-β-D-glucopyranoside 1% wt/vol). Inset: $K_D$ determination of ZnP-DFP3 complex. (ΔAbs followed at 423 nm). (*B*) CD spectrum in the Soret region ZnP-DFP3 at pH 7 (Hepes 50 mM, NaCl 100 mM). (*C*) Electronic spectrum in the visible region of di-$Co^{2+}$-DFP3 complex at pH 7 (Hepes 50 mM, NaCl 100 mM). (*D*) Far-UV CD spectrum of apo-DFP3 (in black), ZnP-DFP3 (in red), di-$Zn^{2+}$-DFP3 (in blue) and di-$Zn^{2+}$-ZnP-DFP3 (in green) at pH 7 (Hepes 50 mM, NaCl 100 mM).

**Di-$Zn^{2+}$-DFP3 Binds to and Stabilizes Radical Semiquinone at Its Binding Site.** We next evaluated the ability of DFP3 to bind to and stabilize a reactive substrate at its dimetal active site. We previously demonstrated the redox-inert di-$Zn^{2+}$ form of a single-domain DF protein is capable of binding to and stabilizing the otherwise highly reactive radical semiquinone form of 3,5-di-tert-butylcatechol (DTBC) (58, 59). In aqueous solution, the semiquinone form of DTBC is much less stable than either the corresponding catechol or quinone (DTBQ). Thus, the semiquinone form is present only in trace amounts when an equimolar mixture of DTBC and DTBQ are mixed in aqueous solution. However, upon addition of stoichiometric di-$Zn^{2+}$-DFP3 or doubly loaded di-$Zn^{2+}$-ZnP-DFP3, the quinone and catechol species slowly con-proportionate to generate the semiquinone, which is then tightly bound to the protein. The accumulation of the semiquinone is clearly visible based on the appearance of strong bands in the near infrared at 740 nm and 800 nm (ε ~ 5,000 $M^{-1}{\cdot}cm^{-1}$) (Fig. 5 *A* and *B*). The resulting radical semiquinone complex showed no signs of decomposition over the course of a day. Thus, DFP3 uses binding energy to stabilize a substrate in a radical semiquinone form that is otherwise reactive and energetically inaccessible at ambient temperature.

**Binding of Porphyrin Allosterically Modulates the Ferroxidase and Oxygenase Activities of DFP3.** Many diiron proteins catalyze the ferroxidase reaction, which involves two-electron oxidation of ferrous ($Fe^{2+}$) to ferric ($Fe^{3+}$) ions with concomitant reduction of molecular oxygen to hydrogen peroxide (39, 60, 61). The kinetics of the reaction can be easily evaluated by monitoring the appearance of the resulting oxo-to-ferric charge-transfer (LMCT) transition near 360 nm. In the presence of molecular oxygen, both apo-DFP3 and ZnP-DFP3 catalyzed the ferroxidase reaction, as evidenced by the formation of LMCT bands, with extinction coefficients 5,000 and ~3,750 $M^{-1}{\cdot}cm^{-1}$ per diiron site for the resulting di-$Fe^{3+}$-DFP3 and di-$Fe^{3+}$-ZnP-DFP3 proteins, respectively (Fig. 5 *C* and *D*). However, the kinetics differed markedly between the two proteins. When apo-DFP3 was incubated with stoichiometric $Fe^{2+}$ ions, the diferric product was fully formed within 1 min (Fig. 5*C*). The slow increase in absorbance at longer times in Fig. 5*C* is likely due to light scattering from aggregation or ferric ion precipitation, which leads to nonspecific increases in absorption at wavelengths well outside the region LMCT region. By contrast, a much slower process was observed when ZnP-bound DFP3 was incubated with ferrous ions under identical conditions, and conversion to the diferric form was not fully complete after 30 min (Fig. 5*D*). Thus, ZnP acts as an allosteric regulator of the dimetal catalytic site of DFP3, having a profound influence on the reactivity of the diferrous state toward molecular oxygen as well as the long-term stability of the bound diferric state of the cofactor.

We next evaluated the ability of DFP3 to catalyze the two-electron oxidation of 4-amino-phenol (4AP) to the corresponding

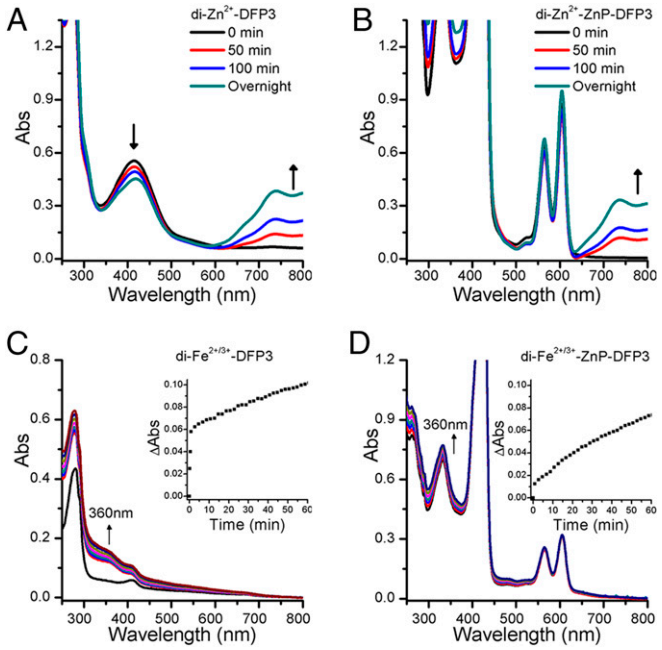



**Fig. 5.** DFP3 simultaneously binds the metal cofactors and exogenous ligands. UV/vis spectral changes upon addition of DTBC and DTBQ mixture, to generate DTBSQ$^{\bullet}$ in situ, to a solution of: di-$Zn^{2+}$-DFP3 (*A*); di-$Zn^{2+}$-ZnP-DFP3 (*B*). (*C*) Ferroxidase activity of apo-DFP3 at pH 7 (Hepes 50 mM, NaCl 100 mM). (*D*) Ferroxidase activity of ZnP-DFP3 at pH 7 (Hepes 50 mM, NaCl 100 mM). In *C* and *D*, the spectra display increase in the absorbance of the typical oxo-$Fe^{3+}$ LMCT band near 360 nm. Inset: time course of the ferroxidase reaction followed at 360 nm.

quinone imine, as studied previously for DF proteins (36–38). The catalytic cycle (Scheme 1) involves the oxidation of the diferrous protein to a diferric species by $O_2$ (the ferroxidase reaction, Scheme 1*A*). The diferric protein then binds and performs a two-electron oxidation of the substrate 4AP to produce the corresponding quinone monoimine (**1** and **2**, in Scheme 1*B*, respectively). The produced diferrous form is then oxidized by $O_2$, thereby initiating another catalytic cycle. To allow easy detection of the quinone imine, it is rapidly quenched with 1,3-diaminobenzene (MPD) to form an aminoindoaniline dye (**3** and **4,** in Scheme 1*C*), with $\lambda_{max}$ at 486 nm at pH 7 in a reaction first studied in 1879 by Witt (62–64).

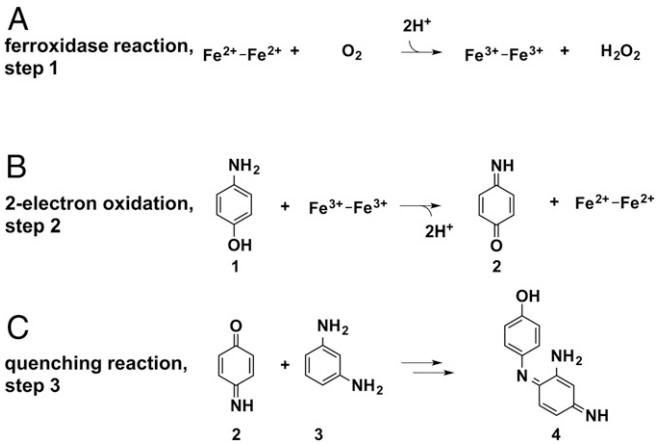
The diiron forms of both di-Fe-DFP3 and di-Fe-ZnP-DFP3 showed significant phenol oxidase activity, which was strongly modulated by the binding of ZnP (Fig. 6 *A*–*C*). Di-Fe-DFP3 displayed saturation kinetics (Fig. 6*D* and *SI Appendix*, Fig. S15), and a nonlinear least squares fit to the data provided Michaelis–Menten parameters: $K_m = 2.9 \pm 0.3$ mM; $k_{cat} = 0.70 \pm 0.04$ min$^{-1}$; $k_{cat}/K_m = 2.4\ 10^2$ M$^{-1}$·min$^{-1}$. Compared to DF3 ($K_m = 1.97 \pm 0.27$ mM and $k_{cat} = 2.72 \pm 0.19$ min$^{-1}$), a slight increase in $K_m$ and a decrease in $k_{cat}$ were observed (36). The reaction progressed over multiple turnovers, and a total of 10 turnovers were observed at [4AP] = 1 mM and [di-Fe-DFP3] = 20 μM.

The Michaelis–Menten parameters for the corresponding ZnP-bound diiron DFP3 protein were strongly affected by the binding of the allosteric modulator ZnP, resulting in a fourfold tighter $K_m$ ($0.68 \pm 0.15$ mM) and a sevenfold decrease in $k_{cat}$ ($0.106 \pm 0.009$ min$^{-1}$), corresponding to a $k_{cat}/K_m$ of $1.6\ 10^2$ M$^{-1}$·min$^{-1}$.

To provide insight into the allosteric modulation, we conducted 1-μs all-atom molecular dynamic (MD) simulations of di-$Fe^{2+}$-DFP3 and di-$Fe^{2+}$-ZnP-DFP3 in triplicate. The most striking difference in the structural ensembles was a shift in the second-shell ligand, Tyr18 (*SI Appendix*, Figs. S16–S18). In the starting model of di-$Fe^{2+}$-ZnP-DFP3, this buried residue donates a hydrogen bond to the metal ion ligand Glu72. Across the three simulations, Tyr18 largely maintains this major conformation, which allows direct and water-mediated hydrogen bonds to Glu72. In a second minor conformer, the Tyr sidechain transiently points toward the surface. In absence of ZnP, Tyr18 shifts and prefers this alternate sidechain conformation, in which its hydroxyl group forms hydrogen bonds with the backbone carbonyl and surface sidechains of a neighboring helix, suggesting that steric blocking of the rotation by ZnP assists the maintenance of the Tyr18–Glu72 interaction in di-$Fe^{2+}$-ZnP-DFP3. Throughout the simulations, there was little evidence of additional differences in the dynamics or structure of the active site or pore between the apo and holo states (active site rmsd values, pore dimensions, iron-ligand distances). Our use of a fixed-charge atomic force field, while able to capture the shift in Tyr18 orientation, is unlikely to capture a potential change in the energetics of the active site resulting from persistent hydrogen bonding of one of the ligating carboxylate oxygens of Glu72, which may produce the shift in catalytic activity.

## Conclusions

Here, we developed a design method to allow tight coupling of protein domains using multiple linkers. Our strategy differs from previously reported approaches, in which one or two linkers were used to fuse two domains in end-to-end manner or through domain insertion. The design strategy was successful not only in preserving the original structural and functional properties of DF and PS1 proteins, but also in allowing the allosteric communication between the domains to control the catalysis. Indeed, the binding of ZnP into PS1 domain markedly decreased $K_m$ and $k_{cat}$ values in the phenol oxidation at DF domain. Further experiments will be needed to fully describe the molecular mechanism of such allosteric control. Nevertheless, MD simulations suggest that it might be associated with a shift in a second-shell Tyr, which was previously found to stabilize the conformation of one Glu ligand in DF. Changes to second-shell ligands can also have significant effects on the midpoint potential of a metal center, which could be relevant to the rate of catalysis. Indeed, mutation of the second-shell Tyr in DF proteins have been shown to largely effect the reactivity of the diiron center (32, 60, 61). The ferroxidase/oxidase reactions require multiple shifts in carboxylate ligands, as the protein adjusts to changes in hydration and protonation that accompany binding of $Fe^{2+}$ and $O_2$. Additionally, the charge on the two metal ions increases from +4 to +6 as they transition from diferrous to diferric. Accordingly, the binding of the ZnP in DFP3 was expected to strongly modulate the rate of these reactions.




**Scheme 1.** Oxidative coupling reactions involving 4-amino-phenol.

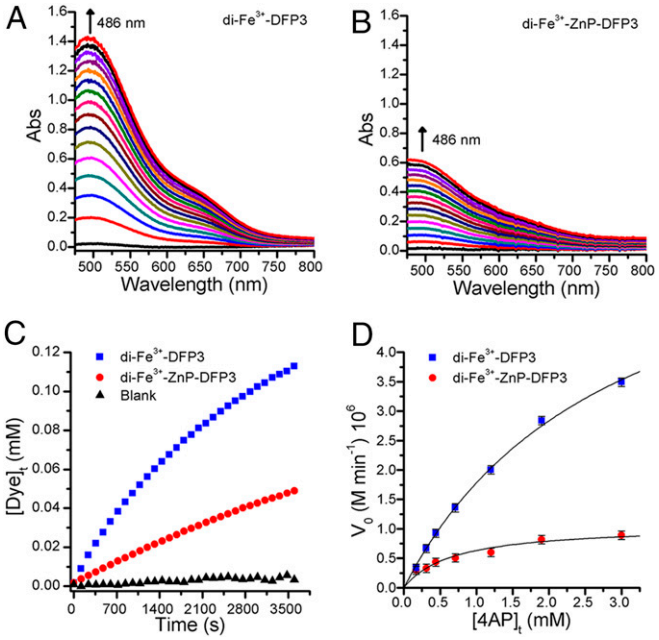



**Fig. 6.** DFP3 shows allosterically regulated phenol oxidase activity. Two-electron oxidation of 4AP in presence of MPD by: 20 μM di-$Fe^{3+}$-DFP3 (*A*); 20 μM di-$Fe^{3+}$-ZnP-DFP3 at pH 7 (Hepes 50 mM, NaCl 100 mM, 10% DMF) (*B*). (*C*) Time courses of 4AP oxidation, by di-$Fe^{3+}$-DFP3 (in blue), di-$Fe^{3+}$-ZnP-DFP3 (in red) and blank (in black), followed at 486 nm. The absorbance values were converted to aminoindoaniline dye concentration. (*D*) Initial rate of the oxidation in function of 4AP concentration catalyzed by di-$Fe^{3+}$-DFP3 (in blue) and di-$Fe^{3+}$-ZnP-DFP3 (in red). Kinetics parameters ($k_{cat}$ and $K_m$) were determined by nonlinear least squares fit of the Michaelis-Menten equation.

In conclusion, these studies provide a prototype for the construction of allosterically regulated proteins and molecular machines.

## Materials and Methods

**Protein Expression, Purification, and Structural Characterization.** Details regarding protein expression, purification, and structural characterization are given in *SI Appendix*. Briefly, genes for the proteins were ordered from IDT, cloned into either a pET-11a or per pET-28a plasmids, and purified via a Ni column, followed by His-tag cleavage by TEV protease.

**SEC.** SEC was carried out for DFP2 and DFP3 on an AKTA FPLC (GE) fitted with a Superdex 75 Increase 10/300 and eluted at 4 °C with Hepes (50 mM, pH 7)/NaCl (100 mM) buffer solution, at a 0.3 mL/min flow rate.

The ZnP–DFP complexes were prepared adding a twofold excess of the porphyrin from a 12 mM dimethyl sulfoxide (DMSO) stock solution to a 50 mM Hepes, 100 mM NaCl, pH 7 buffer with apo-DFP protein (the final DMSO concentrations were kept to <5%). The solution was incubated for 15 min at 70 °C and subsequently filtered before injection. Four size standards were employed for calibration: blue dextran (2,000 kDa), Conalbumin (75 kDa), Carbonic Anhydrase (29 kDa) and Ribonuclease A (13.7 kDa). Standard curve (*SI Appendix*, Fig. S19) and elution parameters (*SI Appendix*, Table S2) are reported. The theoretical Stokes radius values for DFP proteins was calculated by the shell model with hydration layer 3.2 Å of HYDRONMR (65), with parameters corrected according to experimental conditions, starting from the X-ray structure of DFP1.

**Circular Dichroism.** CD measurements were performed using a J-815 spectropolarimeter equipped with a thermostated cell holder (JASCO). CD spectra were collected at 25 °C, from 260 to 200 nm and from 500 and 300 nm at 0.2-nm intervals with a 20 nm·$min^{-1}$ scan speed, at 2.5-nm bandwidth and at 16-s response. Cells of 0.1-cm path length were used in the measurements at protein concentration of 20 μM. Mean residue ellipticities θ were calculated using the equation $\theta = \theta_{obs}/(10 \cdot l \cdot C \cdot n)$, in which $\theta_{obs}$ is the ellipticity measured in millidegrees, *l* is the path length of the cell in centimeters, *C* is the concentration in moles per liter, and *n* is the number of residues in the protein for the UV region and equal to 1 in the Visible region.

**UV-Visible Spectroscopy.** UV-visible (vis) measurements were performed using Cary 60 UV/vis spectrophotometer using quartz cuvettes of 1 cm. Wavelength scans were performed at 25 °C from 200 to 800 nm, with a 60 nm·$min^{-1}$ scan speed.

ZnP (2.2 μM) was solubilized in a 1.5-mL solution of 50 mM Hepes, 100 mM NaCl, pH 7 buffer by inclusion of 1% wt/vol octyl-β-D-glucopyranoside. Four microliters of a 75 μM stock of apo-DFP3 or di-$Zn^{2+}$-DFP3 (0.09 μM aliquots) was titrated into the 1.5-mL solution containing the porphyrin, and an electronic absorption spectrum was measured until >1.5 equivalents of protein were added. Absorbance changes at 423 nm, due to His-Zn coordination-induced spectral shifts of the porphyrin, were fit to a single-site, protein-ligand binding model.

The di-$Co^{2+}$-DFP3 complex was prepared from a stock solution of apo-DFP3 in 50 mM Hepes buffer 100 mM NaCl pH 7.0 by adding 2 μL of a solution of $CoCl_2$ (100 mM in $H_2O$).

The initial DFP3 concentration (500 μM) was spectrophotometrically determined, by using ε(280 nm) = 22.460 $cm^{-1} \cdot M^{-1}$.

Di-$Zn^{2+}$-DFP3 and di-$Zn^{2+}$-ZnP-DFP3 complexes were prepared in the same way as di-$Co^{2+}$-DFP3 complex. However, the final concentration was 50 μM of complex. DTBC and DTBQ, premixed in a separate vessel (in a 1:1 ratio, dissolved in dimethylformamide [DMF]), were added in 10 molar equivalents to the protein solutions and then quickly mixed by pipetting.

To study the ferroxidase reactivity, stock solutions of ferrous ammonium sulfate were prepared by dissolving the solid in 1% sulfuric acid and then diluted to a final concentration of 10 mM (0.05% sulfuric acid). Reaction started upon addition of a stoichiometric amount (2 eq) of $Fe^{2+}$ to a solution of the apo-protein (20 μM) in aerobic conditions.

The oxidation of 4AP by preloaded protein diferric complexes were conducted in the presence of MPD, which led to the formation of the aminoindoaniline dye that allowed for spectrophotometric detection at the absorption maximum of 486 nm. The final solution contained 20 μM protein in 50 mM Hepes buffer 100 mM NaCl pH 7.0 10% DMF, 1 mM 4AP, 10 mM MPD.

The effect of 4AP concentration on catalysis was tested in the range 0.1 mM–3 mM. The experiments were performed always in presence of 10 mM MPD, and 10 μM diferric proteins in 50 mM Hepes buffer 100 mM NaCl pH 7.0 10% DMF. The formation of the aminoindoaniline dye was followed at 528 nm at 25 °C and ε value of 10,700 $M^{-1} \cdot cm^{-1}$, taken from the literature.

**Data Availability.** The atomic coordinates and structure factors have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (https://www.rcsb.org/), under the PDB ID code 7JH6.

**ACKNOWLEDGMENTS.** We acknowledge support from NIH Grant R35 GM12263, NSF Grant 1709506, and NSF Grant CHE-1709497. This material is based on work supported in part by the Air Force Office of Scientific Research under Award FA9550-19-1-0331. We acknowledge support of the University of Naples "Federico II" for the stage of F.P. at University of California, San Francisco, during his PhD.

1. L. Yu *et al.*, Grammar of protein domain architectures. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3636–3645 (2019).
2. R. Aroul-Selvam, T. Hubbard, R. Sasidharan, Domain insertions in protein structures. *J. Mol. Biol.* **338**, 633–641 (2004).
3. K. Forslund, E. L. Sonnhammer, Evolution of protein domain architectures. *Methods Mol. Biol.* **856**, 187–216 (2012).
4. S. Vishwanath, A. G. de Brevern, N. Srinivasan, Same but not alike: Structure, flexibility and energetics of domains in multi-domain proteins are influenced by the presence of other domains. *PLOS Comput. Biol.* **14**, e1006008 (2018).
5. M. Bashton, C. Chothia, The generation of new protein functions by the combination of domains. *Structure* **15**, 85–99 (2007).
6. Y. T. Lai *et al.*, Designing and defining dynamic protein cage nanoassemblies in solution. *Sci. Adv.* **2**, e1501855 (2016).
7. J. F. Matthaei *et al.*, Designing two-dimensional protein arrays through fusion of multimers and interface mutations. *Nano Lett.* **15**, 5235–5239 (2015).
8. V. Stein, K. Alexandrov, Synthetic protein switches: Design principles and applications. *Trends Biotechnol.* **33**, 101–110 (2015).
9. H. Yang, L. Liu, F. Xu, The promises and challenges of fusion constructs in protein biochemistry and enzymology. *Appl. Microbiol. Biotechnol.* **100**, 8273–8281 (2016).
10. K. Kariyawasam *et al.*, An artificial hemoprotein with inducible peroxidase- and monooxygenase-like activities. *Chemistry*, 10.1002/chem.202002434 (2020).

11. X. Chen, J. L. Zaro, W. C. Shen, Fusion protein linkers: Property, design and functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013).
12. T. M. Jacobs *et al*., Design of structurally distinct proteins using strategies inspired by evolution. *Science* **352**, 687–690 (2016).
13. S. J. Youn *et al*., Construction of novel repeat proteins with rigid and predictable structures using a shared helix method. *Sci. Rep.* **7**, 2595 (2017).
14. T. J. Brunette *et al*., Modular repeat protein sculpting using rigid helical junctions. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 8870–8875 (2020).
15. A. Roy, I. Sarrou, M. D. Vaughn, A. V. Astashkin, G. Ghirlanda, De novo design of an artificial bis[4Fe-4S] binding protein. *Biochemistry* **52**, 7586–7594 (2013).
16. G. M. Bender *et al*., De novo design of a single-chain diphenylporphyrin metalloprotein. *J. Am. Chem. Soc.* **129**, 10732–10740 (2007).
17. J. Hong, O. A. Kharenko, M. Y. Ogawa, Incorporating electron-transfer functionality into synthetic metalloproteins from the bottom-up. *Inorg. Chem.* **45**, 9974–9984 (2006).
18. T. A. Farid *et al*., Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).
19. J. A. Mancini *et al*., Multi-step excitation energy transfer engineered in genetic fusions of natural and synthetic light-harvesting proteins. *J. R. Soc. Interface* **14**, 20160896 (2017).
20. R. Kajihara, K. Oohora, T. Hayashi, Photoinduced electron transfer within supramolecular hemoprotein co-assemblies and heterodimers containing Fe and Zn porphyrins. *J. Inorg. Biochem.* **193**, 42–51 (2019).
21. S. H. Mejias *et al*., Repeat proteins as versatile scaffolds for arrays of redox-active FeS clusters. *Chem. Commun. (Camb.)* **55**, 3319–3322 (2019).
22. K. Oohora *et al*., Supramolecular hemoprotein assembly with a periodic structure showing heme-heme exciton coupling. *J. Am. Chem. Soc.* **140**, 10145–10148 (2018).
23. D. V. Zaytsev *et al*., Metal-binding properties and structural characterization of a self-assembled coiled coil: formation of a polynuclear Cd-thiolate cluster. *J. Inorg. Biochem.* **119**, 1–9 (2013).
24. O. A. Kharenko, D. C. Kennedy, B. Demeler, M. J. Maroney, M. Y. Ogawa, Cu(I) luminescence from the tetranuclear Cu4S4 cofactor of a synthetic 4-helix bundle. *J. Am. Chem. Soc.* **127**, 7678–7679 (2005).
25. C. E. Laplaza, R. H. Holm, Helix-loop-helix peptides as scaffolds for the construction of bridged metal assemblies in proteins: The spectroscopic A-cluster structure in carbon monoxide dehydrogenase. *J. Am. Chem. Soc.* **123**, 10255–10264 (2001).
26. B. R. Gibney, S. E. Mulholland, F. Rabanal, P. L. Dutton, Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15041–15046 (1996).
27. J. Grzyb *et al*., De novo design of a non-natural fold for an iron-sulfur protein: Alpha-helical coiled-coil with a four-iron four-sulfur cluster binding site in its central core. *Biochim. Biophys. Acta* **1797**, 406–413 (2010).
28. N. H. Joh *et al*., De novo design of a transmembrane $Zn^{2+}$-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
29. S. Q. Zhang *et al*., De novo design of tetranuclear transition metal clusters stabilized by hydrogen-bonded networks in helical bundles. *J. Am. Chem. Soc.* **140**, 1294–1304 (2018).
30. M. Chino *et al*., Spectroscopic and metal binding properties of a de novo metalloprotein binding a tetrazinc cluster. *Biopolymers* **109**, e23339 (2018).
31. M. Flores *et al*., Copper environment in artificial metalloproteins probed by electron paramagnetic resonance spectroscopy. *J. Phys. Chem. B* **119**, 13825–13833 (2015).
32. A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, De novo design of four-helix bundle metalloproteins: One scaffold, diverse reactivities. *Acc. Chem. Res.* **52**, 1148–1159 (2019).
33. A. Lombardi *et al*., Retrostructural analysis of metalloproteins: Application to the design of a minimal model for diiron proteins. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6298–6305 (2000).
34. M. Chino *et al*., A de novo heterodimeric Due Ferri protein minimizes the release of reactive intermediates in dioxygen-dependent oxidation. *Angew. Chem. Int. Ed. Engl.* **56**, 15580–15583 (2017).
35. M. Chino, L. Leone, O. Maglio, A. Lombardi, Designing covalently linked heterodimeric four-helix bundles. *Methods Enzymol.* **580**, 471–499 (2016).
36. M. Faiella *et al*., An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.* **5**, 882–884 (2009).
37. J. Kaplan, W. F. DeGrado, De novo design of catalytic proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11566–11570 (2004).
38. A. J. Reig *et al*., Alteration of the oxygen-dependent reactivity of de novo Due Ferri proteins. *Nat. Chem.* **4**, 900–906 (2012).
39. P. P. Wei, A. J. Skulan, H. Wade, W. F. DeGrado, E. I. Solomon, Spectroscopic and computational studies of the de novo designed protein DF2t: Correlation to the biferrous active site of ribonucleotide reductase and factors that affect O2 reactivity. *J. Am. Chem. Soc.* **127**, 16098–16106 (2005).
40. O. Maglio *et al*., Artificial di-iron proteins: Solution characterization of four helix bundles containing two distinct types of inter-helical loops. *J. Biol. Inorg. Chem.* **10**, 539–549 (2005).
41. N. F. Polizzi *et al*., De novo design of a hyperstable non-natural protein-ligand complex with sub-Å accuracy. *Nat. Chem.* **9**, 1157–1164 (2017).
42. R. P. Diensthuber, M. Bommer, T. Gleichmann, A. Möglich, Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators. *Structure* **21**, 1127–1136 (2013).
43. N. Koga *et al*., Principles for designing ideal protein structures. *Nature* **491**, 222–227 (2012).
44. K. T. Simons, C. Kooperberg, E. Huang, D. Baker, Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J. Mol. Biol.* **268**, 209–225 (1997).
45. V. Frappier, J. M. Jenson, J. Zhou, G. Grigoryan, A. E. Keating, Tertiary structural motif sequence statistics enable facile prediction and design of peptides that bind anti-apoptotic bfl-1 and mcl-1. *Structure* **27**, 606–617.e5 (2019).
46. C. O. Mackenzie, J. Zhou, G. Grigoryan, Tertiary alphabet for the observable protein structural universe. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7438–E7447 (2016).
47. J. Zhou, A. E. Panaitiu, G. Grigoryan, A general-purpose protein design framework based on mining sequence-structure relationships in known protein structures. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1059–1068 (2020).
48. J. Zhou, G. Grigoryan, Rapid search for tertiary fragments reveals protein sequence-structure relationships. *Protein Sci.* **24**, 508–524 (2015).
49. G. Grigoryan, W. F. Degrado, Probing designability via a generalized model of helical bundle geometry. *J. Mol. Biol.* **405**, 1079–1100 (2011).
50. S. Gentemann *et al*., Variations and temperature dependence of the excited state properties of conformationally and electronically perturbed zinc and free base porphyrins. *J. Phys. Chem. B* **101**, 1247–1254 (1997).
51. S. Hirota, Oligomerization of cytochrome c, myoglobin, and related heme proteins by 3D domain swapping. *J. Inorg. Biochem.* **194**, 170–179 (2019).
52. S. Nagao *et al*., Thermodynamic control of domain swapping by modulating the helical propensity in the hinge region of myoglobin. *Chem. Asian J.* **15**, 1743–1749 (2020).
53. N. Yoshida, M. Higashi, H. Motoki, S. Hirota, Theoretical analysis of the domain-swapped dimerization of cytochrome c: An MD and 3D-RISM approach. *J. Chem. Phys.* **148**, 025102 (2018).
54. T. Miyamoto *et al*., Domain-swapped cytochrome $cb_{562}$ dimer and its nanocage encapsulating a Zn-$SO_4$ cluster in the internal cavity. *Chem. Sci. (Camb.)* **6**, 7336–7342 (2015).
55. S. J. Lahr *et al*., Analysis and design of turns in alpha-helical hairpins. *J. Mol. Biol.* **346**, 1441–1454 (2005).
56. I. Bertini, C. Luchinat, High spin cobalt(II) as a probe for the investigation of metalloproteins. *Adv. Inorg. Biochem.* **6**, 71–111 (1984).
57. R. Torres Martin de Rosales *et al*., Spectroscopic and metal-binding properties of DF3: An artificial protein able to accommodate different metal ions. *J. Biol. Inorg. Chem.* **15**, 717–728 (2010).
58. G. Ulas, T. Lemmin, Y. Wu, G. T. Gassner, W. F. DeGrado, Designed metalloprotein stabilizes a semiquinone radical. *Nat. Chem.* **8**, 354–359 (2016).
59. J. H. Yoon *et al*., Uno Ferro, a de novo designed protein, binds transition metals with high affinity and stabilizes semiquinone radical anion. *Chemistry* **25**, 15252–15256 (2019).
60. R. A. Snyder *et al*., Systematic perturbations of binuclear non-heme iron sites: Structure and dioxygen reactivity of de Novo due Ferri proteins. *Biochemistry* **54**, 4637–4651 (2015).
61. R. A. Snyder, S. E. Butch, A. J. Reig, W. F. DeGrado, E. I. Solomon, Molecular-level insight into the differential oxidase and oxygenase reactivities of de Novo due Ferri proteins. *J. Am. Chem. Soc.* **137**, 9302–9314 (2015).
62. O. N. Witt, Ueber neue farbstoffe. *Ber. Dtsch. Chem. Ges.* **12**, 931–939 (1879).
63. F. Ullmann, J. Gnaedinger, Über indamine. *Ber. Dtsch. Chem. Ges.* **45**, 3437–3446 (1912).
64. J. F. Corbett, Benzoquinone imines. Part V. Mechanism and kinetics of the reaction of p-benzoquinone monoimines with m-phenylenediamines. *J. Chem. Soc. B* **1969**, 823–826 (1969).
65. J. García de la Torre, M. L. Huertas, B. Carrasco, HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**, 138–146 (2000).