



OPEN ACCESS

EDITED BY

Vladimir N. Uversky,
University of South Florida,
United States

REVIEWED BY

Maria Grazia Murralli,
University of California, Los Angeles,
United States
Yi Song,
Downstate Health Sciences University,
United States
Susmit Narayan Chaudhury,
Los Alamos National Laboratory (DOE),
United States

*CORRESPONDENCE

Nathan M. Ennist,
ennisist@uw.edu

SPECIALTY SECTION

This article was submitted to Structural
Biology,
a section of the journal
Frontiers in Molecular Biosciences

RECEIVED 18 July 2022

ACCEPTED 18 August 2022

PUBLISHED 21 September 2022

CITATION

Ennist NM, Stayrook SE, Dutton PL and
Moser CC (2022), Rational design of
photosynthetic reaction center
protein maquettes.
Front. Mol. Biosci. 9:997295.
doi: 10.3389/fmolb.2022.997295

COPYRIGHT

© 2022 Ennist, Stayrook, Dutton and
Moser. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Rational design of photosynthetic reaction center protein maquettes

Nathan M. Ennist^{1,2,3*}, Steven E. Stayrook^{1,4,5}, P. Leslie Dutton¹
and Christopher C. Moser¹

¹Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA, United States, ²Institute for Protein Design, University of Washington, Seattle, WA, United States, ³Department of Biochemistry, University of Washington, Seattle, WA, United States, ⁴Department of Pharmacology, Yale University School of Medicine, New Haven, CT, United States, ⁵Yale Cancer Biology Institute, Yale University West Campus, West Haven, CT, United States

New technologies for efficient solar-to-fuel energy conversion will help facilitate a global shift from dependence on fossil fuels to renewable energy. Nature uses photosynthetic reaction centers to convert photon energy into a cascade of electron-transfer reactions that eventually produce chemical fuel. The design of new reaction centers *de novo* deepens our understanding of photosynthetic charge separation and may one day allow production of biofuels with higher thermodynamic efficiency than natural photosystems. Recently, we described the multi-step electron-transfer activity of a designed reaction center maquette protein (the RC maquette), which can assemble metal ions, tyrosine, a Zn tetrapyrrole, and heme into an electron-transport chain. Here, we detail our modular strategy for rational protein design and show that the intended RC maquette design agrees with crystal structures in various states of assembly. A flexible, dynamic apo-state collapses by design into a more ordered holo-state upon cofactor binding. Crystal structures illustrate the structural transitions upon binding of different cofactors. Spectroscopic assays demonstrate that the RC maquette binds various electron donors, pigments, and electron acceptors with high affinity. We close with a critique of the present RC maquette design and use electron-tunneling theory to envision a path toward a designed RC with a substantially higher thermodynamic efficiency than natural photosystems.

KEYWORDS

***de novo* protein design, electron transfer, heme, porphyrin binding, protein structure, artificial photosynthesis, photochemical charge separation, redox chemistry**

Introduction

The advent of oxygenic photosynthesis *c.* 3 billion years ago has been called the “big bang of evolution” because it precipitated a new age of abundant life under an oxidizing atmosphere (Barber and Tran, 2013; Barber, 2016; Sanchez-Baracaldo and Cardona, 2020). Biology solves the solar energy storage problem by using photosynthetic reaction centers (RCs) to produce oxygen and hydrogen equivalents (fuel) from water splitting.

RCs allow life to tap into the $>10^8$ GW supply of solar power available at the surface of the Earth (World Energy Council, 2013). In contrast, energy storage remains a challenge for solar photovoltaics and other forms of renewable energy, making artificial photosynthesis an important objective for 21st century chemistry.

The RC executes a series of electron-transfer reactions to convert solar photons into chemical energy. The minimalist RC design is a chain of at least three cofactors, one of which is light-activatable, that are connected by electron tunneling through the insulating protein framework. A photoexcited pigment (P) in the middle of the RC reduces an electron acceptor (A) and oxidizes an electron donor (D), creating a charge-separated D^+PA^- state that can be used for redox chemistry such as water oxidation in Photosystem II (PSII). RCs maintain their near-unity quantum yields (Wraight and Clayton, 1974) by tuning the rates of electron tunneling to favor forward electron transfer while preventing electron-hole recombination. To estimate rates of thermodynamically downhill electron-tunneling reactions, the following empirical expression was developed from studies of natural RCs and other redox active proteins (Moser et al., 1992; Page et al., 1999; Moser et al., 2010):

$$\log_{10}(k_{\text{ET}}) = 13 - 0.6(R - 3.6) - 3.1 \frac{(\Delta G + \lambda)^2}{\lambda} \quad (1)$$

Equation 1

The rate constant k_{ET} is given in s^{-1} , the edge-to-edge distance R between redox centers is in Å , and units of eV are used for both the reorganization energy λ and the driving force of the reaction ΔG . For uphill reactions (positive ΔG), the log rate is slower by $\Delta G/0.06$ eV, maintaining a Boltzmann ratio between the downhill and uphill rates (Woodbury et al., 1986).

Several model systems have been developed to study RCs. Synthetic photochemical DPA triads reproduce the multi-step electron-transport activity of native photosystems (Megiatio et al., 2014; Imahori et al., 2001), and various proteins have been engineered for charge separation starting from native or *de novo* scaffolds (Conlan et al., 2009; Hingorani et al., 2014; Kalman et al., 1999; Farid et al., 2013; Soltau et al., 2015; Tebo et al., 2021). However, no previous *de novo*-designed protein has all the elements of a DPA electron-transport chain. A charge-separating RC maquette with substitutable DPA cofactors can serve as a useful testbed to examine the fundamental properties of RCs without the arduous task of isolating and stabilizing native photosystems.

Our long-term goal is to progressively develop a catalytically active *de novo* RC that self-assembles *in vivo* for enhanced biofuel production. Natural photosystems have high activities and efficiencies, but *de novo* RCs could be designed for improved suitability for commercial biofuel production. Natural photosynthesis stores solar energy in ATP and carbohydrates which need to be converted to other fuels before commercial use.

An RC could be engineered to produce a fuel that requires less downstream processing. A new RC powered by near infrared or green light might complement the absorption cross-sections of native photosynthetic complexes such as PSI and PSII which have similar absorbance profiles and therefore compete for the same solar photons (Blankenship et al., 2011). Finally, the photoexcited P680 pigment of PSII has a low enough redox potential to reduce protons or carbon dioxide, meaning that it is possible to split water into an oxidant and a fuel in a single photosystem (Barber and Tran, 2013; Mersch et al., 2015). The one-photosystem model would shorten the extensive photosynthetic electron-transport chain (which loses free energy as electrons flow down an electrochemical potential gradient) and preclude the need for absorption of a second photon by PSI before fuel production.

Recent progress in computational protein design and structure prediction has been rapid and stands to transform our approach to *de novo* protein design in general (Jumper et al., 2021; Baek et al., 2021). The ever-increasing predictive power of new computational methods will continue to refine our understanding of the protein sequence-structure relationship. The role of human intuition will nevertheless remain significant in the design of electron-transfer proteins for some time to come. RC design is more complex than a protein folding problem because of other important considerations such as cofactor binding and specificity, electron-transfer rates, inter-cofactor distances, excited state energies and lifetimes, redox potentials, and reorganization energies. For some protein scaffolds including four-helix bundles, adherence to simple design rules is sufficient to achieve the intended tertiary structure. The choice of a robust, designable helical bundle fold makes it easier to address the many other factors necessary for efficient light-activated charge separation.

We use “maquettes,” scaled-down *de novo*-designed proteins with simple scaffolds, to help tease out design rules for various protein functions. Iterative experimentation and modification of maquette designs allow biochemists to develop their understanding of protein activity and subsequently develop more advanced functions (Lichtenstein et al., 2012; Ennist et al., 2017; Moser et al., 2016; Grayson and Anderson, 2018).

Progression from the first functionless four helix bundle maquettes in the 1980s to light-driven charge separation in the RC maquette illustrates the maquette approach in action. Early work showed that amino acid sequences with a binary pattern of hydrophobic (H) and polar (P) residues in a heptad repeat (i.e. HPPHPPH or HPPHHPP) could be engineered to fold into four-helix bundle proteins (Regan and DeGrado, 1988; Kamtekar et al., 1993). The addition of metal-ligating histidine residues to the hydrophobic core enabled binding of various metalloporphyrins, opening the door to cofactor-mediated electron transfer and light responsive chemistry (Robertson et al., 1994; Rau and Haehnel, 1998; Sharp et al., 1998).

Assembly of both heme B and a Zn porphyrin into the same maquette enabled intra-protein light-driven electron transfer with a short-lived charge separated state (Farid et al., 2013). The four-helix bundle scaffold was also engineered to host a diiron binding site analogous to that of bacterioferritin, ribonucleotide reductase, and other natural oxidoreductases (Summa et al., 1999; Faiella et al., 2009; Torres Martin de Rosales et al., 2010; Lahr et al., 2005; Lombardi et al., 2019). The DeGrado group recently combined a di-metal center and a Zn porphyrin pigment binding pocket into the same scaffold (Pirro et al., 2020). These advances build upon one another to create maquettes with increasing complexity and functionality.

In this article, we detail the design process that produced a recently reported RC maquette (Ennist et al., 2022), investigate its cofactor binding activity, and analyze crystal structures in various stages of assembly. The RC maquette integrates a Co or Fe porphyrin electron acceptor (A-site), a Zn porphyrin pigment (P-site), and an electron donating di-metal center and tyrosine (D-site) for photoactive charge separation. In addition, we will present a vision for the next generation of RC maquettes: we aim to improve the quantum yield and thermodynamic efficiency of charge separation through the addition of low-potential electron acceptors and use of a photooxidative mechanism for polynuclear metal cluster assembly reminiscent of photoactivation in PSII (Bao and Burnap, 2016; Zhang et al., 2017; Vinyard et al., 2019; Cardona, 2016). Using Eq. 1 as a guide, the maquette approach permits the selection and arrangement of redox cofactors to optimize the preservation of chemical redox energy in the photosynthetic light-activated, charge-separated state.

Materials and methods

Protein expression and purification

A gene encoding the RC maquette was purchased from DNA2.0 in a pJexpress414 vector, and the DH maquette gene was purchased from GenScript in a pET-3a vector. Both genes included an N-terminal His₆ tag followed by a TEV protease cleavage sequence. Invitrogen supplied mutagenesis primers, and mutant plasmids were PCR amplified with AccuPrime™ Pfx SuperMix (Invitrogen). Amino acid sequences are given in Supplementary Table S1.

Plasmids were transformed into BL21-(DE3) competent cells (New England Biolabs), and proteins were expressed and purified as previously described (Ennist et al., 2022). Where noted, an additional purification step of high performance liquid chromatography (HPLC) was executed on a Waters reverse-phase HPLC system using a C4 or C18 HPLC column (Grace Davison Discovery Sciences). HPLC-pure samples were lyophilized and resuspended in >6.5 M GdnHCl, refolded by dilution, and purified by

size exclusion chromatography (SEC) with Superdex 75 prep grade gel filtration medium (GE Healthcare Life Sciences).

Ultraviolet/visible spectroscopy and preparation of tetrapyrrole cofactors

Ultraviolet/visible absorbance spectroscopy (UV/vis) was carried out on a Varian Cary-50 spectrophotometer. Cofactor binding titrations were performed at room temperature using 1 ml samples of ~1 μM RC maquette in quartz cuvettes. Protein concentrations were calculated by assuming that Trp and Tyr residues have extinction coefficients at 280 nm ($\epsilon_{280\text{nm}}$) of 5,500 M⁻¹ cm⁻¹ and 1,490 M⁻¹ cm⁻¹, respectively. Samples of Zn 5-phenyl 15-(*p*-carboxyphenyl) porphyrin (ZnPCP) and Zn Newkome porphyrin were provided by Tatiana Esipova from the research group of Sergei Vinogradov (Kodali et al., 2017). The Zn chlorins SE370 and SE375 were produced by the laboratory of Jonathan Lindsey following previously established methods (Aravindu et al., 2013; Kodali et al., 2017). Other tetrapyrroles were purchased from Frontier Scientific, Inc. Stock solutions of water-soluble cofactors were prepared in water, and insoluble tetrapyrroles were prepared in dimethyl sulfoxide stock solutions of 100–200 μM concentration. Heme B stock concentrations were measured by hemochrome assay (Berry and Trumpower, 1987). Stock concentrations of other tetrapyrrole cofactors were estimated by mass measurements.

For each titration, a cofactor stock solution was titrated into the protein solution in the cuvette. After each addition of cofactor, at least 10 min were allowed for equilibration before collecting a UV/vis spectrum. Cofactor was typically added in steps of 0.1 equivalents of cofactor per protein until after the binding capacity of the protein appeared to have been reached, at which point 0.2 equivalents of cofactor per protein were added at a time. Binding titrations continued until at total of at least three equivalents of cofactor per protein had been added. To determine dissociation constants, we used the following one-site binding equation:

$$A = c_{\text{tot}}\epsilon_{\text{free}} + \frac{(\epsilon_{\text{bound}} - \epsilon_{\text{free}})(K_D + P_{\text{tot}} + c_{\text{tot}} - \sqrt{(K_D + P_{\text{tot}} + c_{\text{tot}})^2 - 4P_{\text{tot}}c_{\text{tot}}})}{2} + b \quad (2)$$

In Eq. 2, A represents the total absorbance at a given wavelength, c_{tot} is the total cofactor concentration, P_{tot} is the total protein concentration, K_D is the dissociation constant, ϵ_{free} is the extinction coefficient of unbound cofactor in solution, ϵ_{bound} is the extinction coefficient of protein-bound cofactor, and b is the absorbance of the sample when c_{tot} is zero.

Estimates of the amounts of heme B incorporated into proteins during overexpression were made by UV/vis measurements. The extinction coefficient of heme B at the Soret maximum ($\epsilon_{415\text{nm}}$) was found to be 115,000 M⁻¹ cm⁻¹

for hemes in bis-His binding sites in the RC maquette and DH maquette. The $\epsilon_{280\text{nm}}$ of heme B bound to a maquette is approximately $20,000 \text{ M}^{-1} \text{ cm}^{-1}$, allowing estimation of the ratio of heme to maquette protein after purification.

Circular dichroism spectroscopy

An Aviv Model 410 instrument was used for circular dichroism (CD). Ultraviolet CD spectra were collected from 190 to 260 nm. The bandwidth used was 1 nm, and data points were measured every 1 nm with 5 s averaging time per data point. Samples were prepared in 0.1 cm path length quartz cuvettes. Each reported spectrum represents the average of three spectra measured in succession. Thermal denaturation experiments measured the CD signal at 222 nm for 5 s with a 2 nm bandwidth, and data were collected every 2°C with a heating rate of 10°C/min followed by a 4-min incubation time. The mean molar ellipticity per residue, θ_{MMR} , is the measured ellipticity divided by 196 residues (the length of the RC maquette).

Crystallography

We report two new crystal structures deposited in the PDB under accession codes 8D9O and 8D9P. The 8D9O protein stock solution was purified by HPLC to remove residual heme B leftover from expression in *E. coli*. Heme B was added to the 8D9P protein sample as described previously (Ennist et al., 2022). Protein stock solutions contained 20–40 mM NaCl and 10–20 mM piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES) buffer at pH 6.5. Crystals were grown in hanging drops in which 1 μl of well solution was added to 1 μl of protein stock solution. Supplementary Table S2 gives protein stock concentrations, well solutions, cofactors, and cryoprotectants used for the two crystals. Both X-ray diffraction data sets were obtained from single crystals grown at 4°C and cooled to 100 K for data collection. A Rigaku Micromax-007 HF rotating copper anode X-ray generator with VariMax HF optics and a Rigaku Saturn 944 HG CCD detector collected X-ray diffraction data. Phaser (McCoy et al., 2007) performed molecular replacement using the crystal structure with PDB ID: 5VJS as a search model (Ennist et al., 2022). Structure solution and refinement was carried out using the CCP4i (Winn et al., 2011) and PHENIX (Adams et al., 2010; Liebschner et al., 2019) software packages. Scaling of diffraction intensities was done in SCALA (Evans, 2006). Structures were refined using REFMAC5 (Vagin et al., 2004), phenix.refine (Afonine et al., 2012), and PDB_REDO (Joosten et al., 2014). RESOLVE performed density modification (Terwilliger, 2000), and Coot facilitated real space refinement (Emsley et al., 2010). Polder omit maps in Supplementary Figure S2 were generated in Phenix (Liebschner et al., 2017). Data collection and refinement statistics are given in Supplementary Table S3.

Electron-tunneling simulations

Equation 1 is used to calculate the electron-tunneling rate for exergonic electron transfers. The reverse, endergonic reaction applies a rate penalty of one order of magnitude for every 0.06 eV of endergonic free energy. These rates are used to generate a set of differential equations solved in Wolfram Mathematica, Version 11.1.1 for electron transfers between cofactors in the tetrad and pentad kinetic schemes. The time zero boundary condition begins with the singlet excited pigment. Supplementary Table S5 provides the inter-cofactor edge-to-edge distances, excited state energies and decay lifetimes, reorganization energies, and redox midpoint potentials used in the calculations. Driving forces for electron transfer are estimated from the midpoint potentials given in Supplementary Table S5.

Computational modeling of protein structure

Comparisons between protein structures were made by aligning the structures and calculating the root mean square deviations between aligned C_{α} atoms (C_{α} RMSD) using TM-align (Version 20190822) (Zhang and Skolnick, 2005). Structure prediction with RoseTTAFold was done using the server at <https://rosetta.bakerlab.org> (RoseTTAFold option) (Baek et al., 2021). Structure prediction with AlphaFold2 (Jumper et al., 2021) was done in ColabFold without multiple sequence alignments (Mirdita et al., 2022). The impact of charge complementarity in the RC maquette as illustrated in Figure 1 was evaluated for two mutant designs containing six point mutations each. The first mutant switches the *c*-position charged amino acids in the porphyrin binding domain in helices 1 and 3: E8K, E15K, E22K, K109E, K116E, and K123E. The second mutant switches *c*-positions of helices 2 and 4: E80K, E87K, E94K, K177E, K184E, and K191E.

Results

Modular design of a reaction center maquette

We conceived the RC maquette as a four-helix coiled coil protein for two primary reasons. First, helical bundles are highly designable. Decades of studies have refined the design rules and shown how diverse amino acid sequences can adopt this fold (Regan and DeGrado, 1988; Harbury et al., 1993; Robertson et al., 1994; Kamtekar et al., 1993; Kamtekar and Hecht, 1995; Moser et al., 2016; Grigoryan and DeGrado, 2011; Rojas et al., 1997; Gibney et al., 1999; Huang et al., 2014; Boyken et al., 2016; Walshaw and Woolfson, 2001; Kumar and

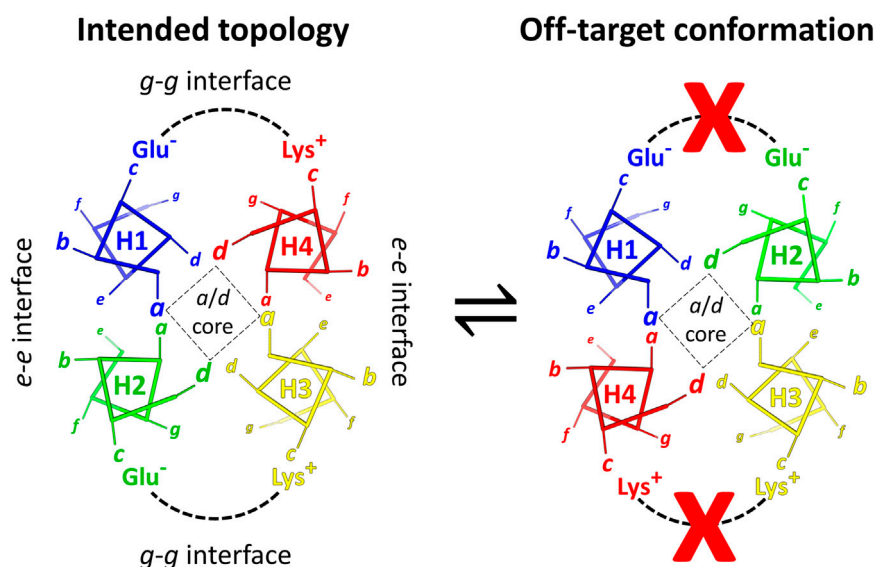


FIGURE 1

Helical wheel diagrams show how charge pairing interactions disfavor a misfolded state. Left: the correctly-folded antiparallel coiled coil structure has favorable Glu – Lys interactions. Right: helices H2 and H4 are transposed in the misfolded state, creating Glu – Glu and Lys – Lys charge clashes as indicated by the red Xs. For each helix, one example heptad is displayed. The distance of an amino acid from the viewer is represented by the size of its heptad letter designation. In helices, thick sticks connect C_{α} atoms, while thinner lines represent C_{α} – C_{β} bonds.

Woolfson, 2021). Approximately 5% of known natural protein sequences contain coiled coils (Rackham et al., 2010), including some that serve as excellent models for the RC maquette. Bacterioferritin, for example, contains a diiron site with a nearby tyrosine amino acid (Conlan et al., 2009; Hingorani et al., 2014), and cytochromes (cyt.) b and b_6 (from the cyt. bc_1 and b_6f complexes, respectively) assemble two porphyrins within electron-transfer distance of one another (Malone et al., 2019). Secondly, a four-helix bundle is an ideal scaffold because it is conducive to the iterative redesign of an RC maquette. Inter-cofactor distances can be adjusted by translating binding sites in increments of approximately one helical turn along the bundle axis. Since cofactor binding sites are largely independent, modules can be added or removed. Simply lengthening the helices will extend the bundle and allow it to assemble more redox centers into a longer chain. Analogous manipulations in many other protein folds would be prohibitively difficult.

Several key principles of four-helix bundle design derived from previous studies contributed to the RC maquette design process. They include 1) binary patterning in heptad repeats, 2) a balance of high helix-propensity vs. beta-branching/aromatic residues, 3) negative design using charge complementarity, 4) shape complementarity in the hydrophobic core, and 5) attention to the structural details of analogous native and designed proteins. We elaborate on these five points below.

Binary patterning

The RC maquette employs a binary pattern of hydrophobic and polar amino acids in heptad repeats throughout its helices to help stabilize a coiled coil fold (Kamtekar et al., 1993). In contrast to straight α -helices that have about 3.6 amino acids per helical turn, canonical coiled coils wrap α -helices around each other like a rope to create a left-handed supercoil with almost exactly 3.5 residues per turn (Grigoryan and Degrado, 2011; Crick, 1953b; Crick, 1953a). As a result, there are precisely two turns per heptad of amino acids, and every seventh residue finds itself in a nearly equivalent local environment. Following convention, we designate the amino acid positions in the RC maquette heptads a through g , such that a - and d -positions are the most buried in the hydrophobic core. With exceptions for cofactor ligation, the RC maquette uses hydrophobic residues in the buried a - and d -positions and polar residues in exposed b -, c -, and f -positions. Hydrophobicity of interfacial e - and g -positions vary along the helices according to the needs of nearby cofactor binding sites.

Amino acid composition

To promote the folding of a uniquely-structured, stable coiled coil, the RC maquette hosts a balance of different types of amino acids in buried a -, d -, e -, and g -positions. Amino acids with high α -helical propensities, such as Ala, Met, and Leu, have a strong preference for α -helical phi/psi angles (Pace and Scholtz, 1998), but an excess of small or flexible residues often confers a poorly ordered and non-native-like

hydrophobic core. Inflexible aromatic and β -branching residues, including Phe, Tyr, Ile, Thr, and Val, help to limit the conformational freedom of the core and constrain it to a single unique structure (Betz and DeGrado, 1996; Raleigh and Degrado, 1992; Gibney et al., 1999; Kamtekar and Hecht, 1995). Helical segments of the RC maquette include 44 Ala and Leu residues (22% of the protein) to promote helix formation and 20 aromatic and β -branching residues (10% of the protein, not including His) to minimize conformational freedom.

Negative design

Negative design using charge complementarity can stabilize the correctly folded state relative to off-target conformations. In Figure 1, a diagram of the RC maquette on the left shows helices 1–4 arranged counterclockwise when viewed from the electron accepting end, but an unintended clockwise topology shown on the right could also form a canonical antiparallel coiled coil. This off-target conformation would have a similar overall structure with favorable core packing and knob-into-hole interactions, but the metal binding site would be distorted. One strategy to favor the intended topology is to create an energy gap between the two alternative states using strategic placement of charged amino acids (Cochran et al., 2005; Summa et al., 2002; Marsh and DeGrado, 2002). In the RC maquette, surface-exposed *c*-positions in the pigment/electron acceptor region were chosen to be charged Glu or Lys such that they would form stabilizing salt bridges in the designed counterclockwise topology and destabilizing charge clashes in the clockwise state (Figure 1). Uncharged Gln residues were selected for most *b*-, *e*-, and *f*-positions in the porphyrin domain to avoid electrostatic interference.

The impact of charge pairing was evaluated by structure prediction using AlphaFold2 in CoLabFold (Jumper et al., 2021; Mirdita et al., 2022). AlphaFold2 accurately predicts the experimentally-validated counterclockwise arrangement of helices in the RC maquette with high confidence (consistent with Figure 1, left). However, structure prediction becomes more ambiguous when six substitution mutations rearrange the distribution of *c*-position Glu and Lys residues of the porphyrin-binding domain to favor the reversed (clockwise) helix threading. Structures of the mutants are predicted with lower confidence, and AlphaFold2 produces some models with the reversed topology. (See methods for details). This illustrates the utility of inter-helical electrostatic interactions in determining tertiary structure.

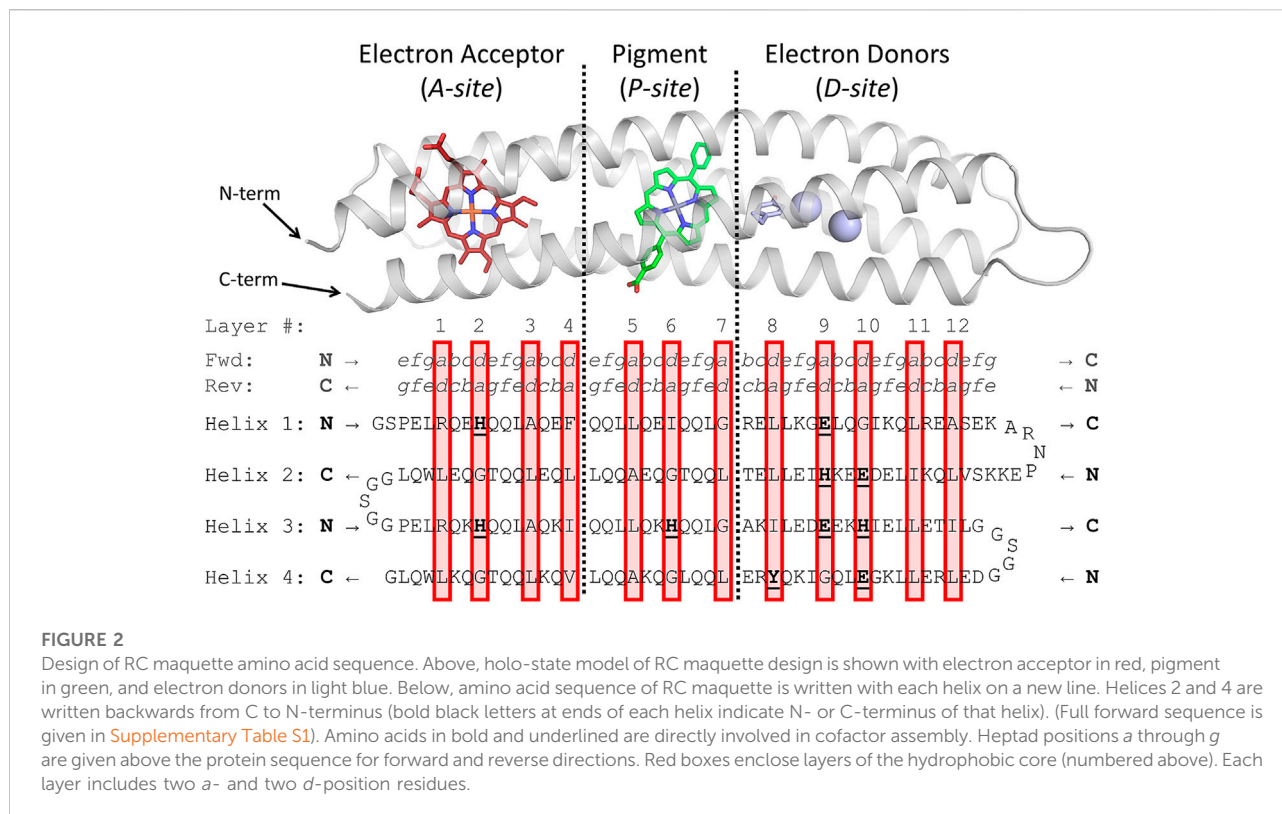
Hydrophobic core packing

The primary driving force for folding of most proteins is the hydrophobic effect, but complementary packing of hydrophobic residues that are far apart in sequence space can be difficult to

predict. Early maquettes solved the problem by using Leu residues almost exclusively in *a*- and *d*-positions (Regan and DeGrado, 1988; Robertson et al., 1994); the flexible side chains with uniform size ensured that the folded state would have a hydrophobic core without destabilizing cavities. However, Leu flexibility also allowed multiple conformational states to exist simultaneously, which complicates characterization efforts (Raleigh and Degrado, 1992; Gibney et al., 1999). Today, computational methods such as AlphaFold (Jumper et al., 2021), RoseTTAFold (Baek et al., 2021), and Rosetta (Leman et al., 2020) can meet the challenge of core packing, but an intuitive, logical approach to sequence selection affords the designer a firmer grasp of the role of each amino acid. A deeper understanding of the protein structure and the reasoning behind the sequence facilitates iterative redesign. One approach is to consider the hydrophobic core of the helical bundle as a stack of layers, where each of the four helices contributes one amino acid to each layer (Munson et al., 1994; Betz and DeGrado, 1996; Betz et al., 1997; Liu et al., 2007; Gernert et al., 1995; Chino et al., 2015; Willis et al., 2000; Summa et al., 1999; Walshaw and Woolfson, 2001; Glover et al., 2014). In the antiparallel RC maquette coiled coil, each layer of the core consists of two *a*- and two *d*-position residues (Figure 1). The layers, indicated by red boxes in Figure 2, are designed to contain a roughly consistent cumulative volume of amino acid side chains and cofactors to avoid introducing cavities or bulges that would destabilize the core. For instance, layer 4 in Figure 2 has large Phe, Ile, Val, and Leu residues, but layer 2 offsets the bulk of its porphyrin electron acceptor ligated by two *d*-position His residues with small *a*-position Gly residues. Not all coiled coils have unambiguously identifiable layers, but in many cases layers help the protein designer recognize and understand patterns of core packing.

Using natural proteins as templates

Distinct *a/d* layers are observable in native proteins including cyt. *b* and *b*₆, as illustrated in Supplementary Figure S1 (Hunte et al., 2000; Malone et al., 2019). Despite being membrane proteins, these two-heme cytochromes serve as useful templates for the design of the pigment/acceptor module of the RC maquette. Their edge-to-edge distance between hemes is close to the optimal *P*-to-*A* distance for long-lived charge separation in the RC maquette (Ennist et al., 2022). Sequence alignments reveal that the *a/d* layer between the hemes of cyt. *b* and *b*₆ (corresponding to layer 4 in the RC maquette; see Figure 2) typically consists of bulky hydrophobic residues that serve to widen the helical bundle so that it can accept the heme cofactors (Rau and Haehnel, 1998; Esposti et al., 1993). Adjacent to this central layer are two heme-contacting layers (corresponding to layers 3 and 5 in Figure 2). They usually include smaller or more flexible amino acids such as Ala and Leu which make space for the heme. The heme-ligating layers (corresponding



to layers 2 and 6 in the RC maquette) use *d*-position His to coordinate the heme iron and small *a*-position “notch glycines” that offset the bulk of the hemes and His residues (Saribas et al., 1997). While Gly residues are generally more destabilizing in water soluble proteins than in membrane helices, studies have shown that protein maquettes can be designed to support multiple helical glycines and remain stable (Torres Martin de Rosales et al., 2010; Faiella et al., 2009; Ghirlanda et al., 2004; Rau and Haehnel, 1998). The RC maquette includes a total of 10 α -helical glycine residues that were selected to optimize core packing and cofactor binding. The sequence identity between the RC maquette and natural membrane *b*-type cytochromes is negligible, but attention to the packing of layers in the hydrophobic core helps give the RC maquette a similar overall structure in the porphyrin-binding module.

Antiparallel four-helix bundles that bind porphyrins are common in nature (Hunte et al., 2000; Malone et al., 2019; Mathews et al., 1979; Weber et al., 1980; Yankovskaya et al., 2003; Dautant et al., 1998), and experimental structures of *de novo*-designed versions have been published recently (Polizzi et al., 2017; Pirro et al., 2020; Mann et al., 2021; Hutchins et al., 2020). One common trend in all of the examples cited here is that the plane of the porphyrin ring cuts through the

e-e interface of the helical bundle, causing the *e-e* interface to expand while the *g-g* interface can remain narrow. This preference may be the result of structural differences between the two interfaces; in *g*-positions, the C_{α} -to- C_{β} bond points toward the hydrophobic core, whereas *e*-position C_{α} -to- C_{β} bonds tend to direct the side chains away from the core, enabling *e*-position residues to more easily avoid clashing with the porphyrin cofactor. Both *e*- and *g*-positions are interfacial and can be either polar or nonpolar in four-helix bundles. We chose mostly polar Gln for *e*-positions in the porphyrin-binding module, and *g*-positions are mostly Leu except for three *g*-position Thr that make second shell hydrogen bonds to His ligands.

An electron donating metal site and tyrosine are found in native carboxylate-bridged diiron four-helix bundle proteins such as bacterioferritin. The DeGrado and Lombardi groups reproduced this distinctive fold in a series of *de novo*-designed “Due Ferri” maquettes and engineered a variety of functions including ferroxidase and phenol oxidase activity (Summa et al., 1999; Faiella et al., 2009; Torres Martin de Rosales et al., 2010; Lahr et al., 2005; Pirro et al., 2020; Chino et al., 2015; Lombardi et al., 2019). Given the success of this line of research, Due Ferri 3 (DF3) was selected as the basis for sequence design in the electron donor module of the RC maquette (Faiella et al., 2009; Torres

Martin de Rosales et al., 2010). DF3 has a higher affinity for various metal ions than previous Due Ferri proteins owing to its four active site Gly residues that impart greater flexibility to the metal-coordinating amino acids (Torres Martin de Rosales et al., 2010). To avoid potential side reactions in the RC maquette, Tyr and Trp amino acids were replaced by redox-inactive residues, with a single Tyr kept as an intermediate electron donor between the metal center and pigment binding site. To avoid complications from nonspecific cofactor ligation, a DF3 loop containing a His residue was replaced with a loop from DF2t (Lahr et al., 2005). A flexible Gly-rich linker was added to incorporate the homodimeric Due Ferri sequence into the single chain RC maquette (see Figure 2).

To fuse the electron donating di-metal/Tyr portion of the RC maquette to the pigment/electron acceptor domain, we joined the corresponding helices together to make an extended 7-nm long antiparallel four-helix bundle. Helix-helix interface widths differ between the two domains. The di-metal/Tyr region has a narrow *e-e* interface that must be widened in the pigment/acceptor domain in order to allow binding of the tetrapyrroles. To facilitate variations in the coiled coil Crick parameters across the RC maquette, each helix has a buried Gly residue at the junction between the two domains to increase backbone flexibility. These four Gly residues also create space for binding of the Zn porphyrin pigment (similar to notch glycines in cyt. *b* and *b*₆) and prevent hydrophobic collapse of the pigment binding site in the apo-state. Independently, the DeGrado group joined two maquettes to make a slightly shorter helical bundle containing one Zn porphyrin and a di-metal site for a different purpose; the Zn porphyrin acted as an allosteric regulator of ferroxidase and oxidase activity in the diiron site (Pirro et al., 2020). In the RC maquette, the addition of an electron accepting porphyrin enables the redox centers to function collectively as a light-driven electron-transport chain.

Figure 2 provides the overall design architecture of the RC maquette and the amino acid sequence. In all, the RC maquette is 196 residues long with a molecular weight of 22.5 kDa and a theoretical isoelectric point of 5.5. The results presented below give support to the design strategy outlined in the present section by providing experimental evidence of cofactor binding activity, a demonstration of modularity in the design, and high resolution X-ray crystal structures.

Cofactor binding in the RC maquette

Binding of heme and other electron acceptors

Expression of the RC maquette in *E. coli* produced red cell pellets suggestive of high-affinity *in vivo* heme binding. The RC maquette retained its red color through Ni-NTA and SEC column chromatography. After purification, ~0.1 equivalents

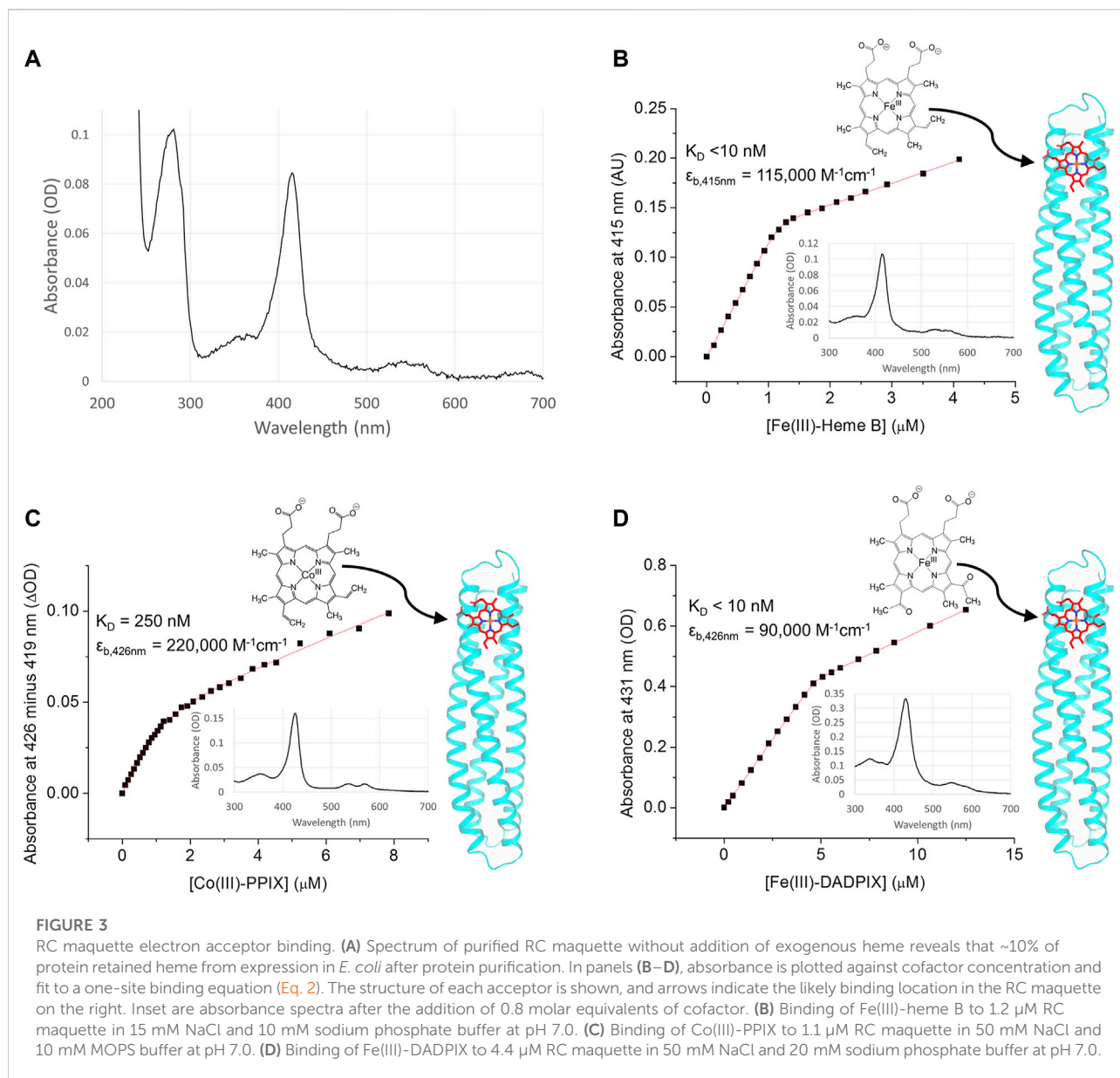
of ferric heme B per mole of RC maquette remained bound, as estimated from UV/vis spectra (Figure 3A). HPLC removed heme from RC maquette samples prior to binding assays.

The RC maquette binds iron tetrapyrroles and cobalt tetrapyrroles through the use of two axial histidine ligands (His9 and His110) and hydrophobic contacts flexible enough to accommodate different alternative cofactors. Binding titrations monitored by UV/vis spectroscopy take advantage of the pronounced changes in metalloporphyrin spectra upon His ligation to quantify RC maquette affinities for various electron acceptors. Figure 3B plots the absorbance of a solution of 1.2 μM RC maquette at 415 nm against ferric heme B concentration. Fitting to Eq. 2 yields a dissociation constant (K_D) of <10 nM and shows that heme binds with the intended stoichiometry of 1.0 heme per maquette. The related iron porphyrin Fe(III) 2,4-diacetyl deuteroporphyrin IX (Fe(III)-DADPIX) binds with a comparable affinity and stoichiometry. Replacement of the ferric ion of heme B with a cobaltic ion gives Co(III) protoporphyrin IX (CoPPIX), an electron acceptor that has been shown to be capable of hydrogen evolution upon reduction (Sommer et al., 2016). Like heme B, CoPPIX accepts two axial His ligands. A spectroscopic titration shows that the RC maquette binds 1.0 CoPPIX per protein with a K_D of 250 nM (Figure 3C). This is about 25 times weaker than the heme B affinity of the RC maquette, but comparable to the CoPPIX affinity of the native hemoprotein cyt. *b*₅₆₂, which has a K_D in the range of 8.9 μM (Sommer et al., 2016) to 45 nM (Alcala-Torano et al., 2021). Despite the presence of His residues at other sites in the RC maquette, none of the electron acceptors tested showed signs of binding outside the designed acceptor binding site. The roles of His9 and His110 in heme coordination were verified by X-ray crystallography (*vide infra*).

Pigment binding

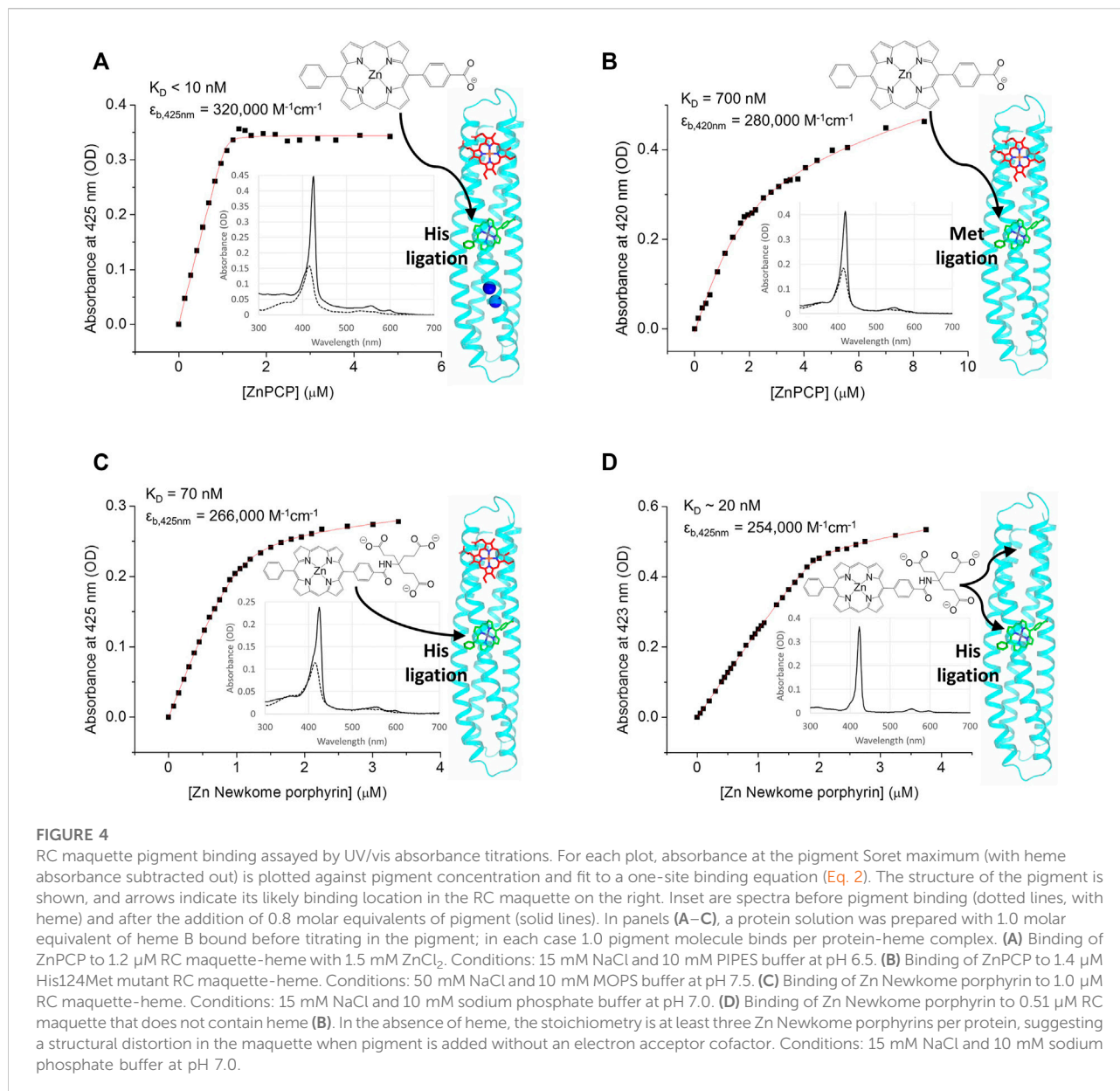
The RC maquette is designed to bind a light-activatable tetrapyrrole pigment in its central pigment binding site, the *P*-site, through the use of hydrophobic contacts and ligation by His124. To test for pigment binding, RC maquette samples were prepared with heme B already bound, and Zn tetrapyrrole binding was monitored spectroscopically as the pigments were titrated into the protein-heme solutions (Figures 4A–C). Zn 5-phenyl,15-(*p*-carboxyphenyl) porphyrin (ZnPcP) exhibits a pronounced red shift in its Soret maximum from 409 nm in buffer to 422 nm upon His ligation with a further shift to 425 nm in the presence of metal dications including Mn(II), Fe(II), and Zn(II) (Ennist et al., 2022). Fitting to Eq. 2 reveals that ZnPcP binds with the expected stoichiometry of 1.0 and a K_D of <10 nM in the presence of Zn(II).

Substitution of methionine (Met) for His124 weakens the ZnPcP affinity to 700 nM with a Soret maximum at 420 nm (Figure 4B). The alteration of ZnPcP absorbance and binding kinetics in the His124Met mutant provides evidence that residue 124 coordinates the Zn ion. Met ligation of Zn



chlorin e_6 has been confirmed in a modified bacterioferritin (Conlan et al., 2009; Hingorani et al., 2014), but to our knowledge, Met ligation of Zn tetrapyrroles has not been tested in maquettes previously. The more hydrophobic Met ligand might be expected to raise the pigment midpoint potential relative to His-ligated pigments as it does for heme B in other maquettes (Moser et al., 2016), offering a possible strategy to fine tune driving forces for electron-transfer reactions. Transient spectroscopy results reported previously indicate an increased quantum yield of charge separation in a Met124 RC maquette variant relative to His124 constructs (Ennist et al., 2022).

Several different Zn tetrapyrrole pigments were tested for binding to the RC maquette (with His124) with heme B bound to the acceptor site. Two Zn diphenyl porphyrins and two Zn diphenyl chlorins bind stoichiometrically with K_D s in the nM range, and several other Zn porphyrins display weak (μM -range K_D) or no detectable affinity. Comparison of the strong- and weak-binding pigments in Supplementary Table S4 reveals a pattern of structural constraints that can be used to predict affinities. As observed previously in other maquettes, high-affinity tetrapyrrole cofactors tend to have amphiphilic character that allows partitioning into the hydrophobic binding site of the protein without loss of



aqueous solubility (Solomon et al., 2014; Kodali et al., 2017). However, in the pigment binding site of the RC maquette, the shape of the cavity confers additional selectivity. Consideration of spectroscopic and crystallographic data suggests that Zn meso-tetraphenyl porphyrins are too bulky to fit into the pigment site. Similar to previous porphyrin-binding maquettes with hydrophobic binding sites (Fry et al., 2010; Cochran et al., 2005; Bender et al., 2007), Zn protoporphyrin IX and Zn chlorin *e*₆ have weaker affinities likely due to carboxylate groups close to the tetrapyrrole ring that resist burial. The four high-affinity Zn tetrapyrroles have a shape that better conforms to the

contours of the binding cavity, and their polar groups are more distant from the central Zn ion and do not need to become buried upon binding.

Owing to the structural similarity between Zn tetrapyrrole pigments and the Fe and Co tetrapyrrole electron acceptors, some of the Zn tetrapyrroles tested were able to bind to both the pigment site and acceptor site (Figure 4D). For this reason, it was necessary to add the electron acceptor first in order to block pigment access to the acceptor site. While acceptor site promiscuity enables testing of different cofactors with varying redox potentials for electron-transfer studies (Ennist et al., 2022), future *in*

vivo self-assembly of a working RC maquette will require greater binding selectivity in the acceptor site to avoid improper assembly of the electron-transport chain.

Metal binding

Carboxylate-bridged di-metal maquettes are typically capable of binding a variety of transition metals (Torres Martin de Rosales et al., 2010). For the purposes of an RC maquette, the most interesting metals are redox-active cations such as Mn(II), Fe(II), and Co(II). Binding of these dications was confirmed spectroscopically in the presence and absence of porphyrin cofactors as described in (Ennist et al., 2022). To summarize, Mn(II) binding was shown to increase the thermostability of the RC maquette, as measured by circular dichroism. Co(II) exhibited an absorbance maximum at 545 nm with an extinction coefficient of $92 \text{ M}^{-1}\text{cm}^{-1}$ when bound to the RC maquette, consistent with 5-coordinate Co(II) binding as expected (Bertini and Luchinat, 1984; Torres Martin de Rosales et al., 2010). The RC maquette was shown to have ferroxidase activity typical of designed and native diiron proteins (Yang et al., 2000; Pasternak et al., 2001). When ZnPCP and heme B are present, Mn(II) binding elicited a 3-nm shift in the Soret band that is reversible upon addition of EDTA. These results illustrated the intended metal binding behavior in the RC maquette (Ennist et al., 2022).

Porphyrin-only DH maquette provides evidence of design modularity

The four-helix coiled coil scaffold facilitates modular design; amino acid selection tailored for specific cofactor binding at one site is largely independent of other sites. The RC maquette is composed of a cyt. *b/b₆*-like pigment/electron acceptor (*P/A*) module and a bacterioferritin-like electron donor (*D*) module. The *D* module is based on the Due Ferri family of homodimeric maquettes, which function independently of a *P/A* module (Faiella et al., 2009; Torres Martin de Rosales et al., 2010; Lombardi et al., 2019; Chino et al., 2015). In contrast, the *P/A* module was designed from scratch with guidance from studies of natural and designed porphyrin-binding proteins as described above.

To examine the *P/A* module separate from the *D* module and to illustrate the validity of our modular approach to design, we developed the Di-Heme maquette (DH) as an independent version of the *P/A* module of the RC maquette (Figures 5, 6). The DH maquette conserves 96% of the sequence of the *P/A* module but inserts loops and N-terminal caps to account for the absence of the *D* module. In addition, a His and second-shell Thr were substituted into the pigment binding site to convert it into a second heme site. While sequence-based structure prediction using AlphaFold (Jumper et al., 2021) or RoseTTAFold (Baek et al., 2021) was not available when the DH maquette was designed, retrospective structure

prediction of DH from its sequence produces apo-state models in good agreement with the *P/A* module of RC maquette crystal structures ($<2.0 \text{ \AA } C_{\alpha}$ RMSD).

IPTG-induced overexpression of the DH maquette in *E. coli* produced red cell pellets similar to the RC maquette. After purification by immobilized metal affinity chromatography, cleavage of the N-terminal His-tag using TEV protease, and SEC, UV/vis spectra showed that the DH maquette contained ~ 0.25 molar equivalents of heme B (or ~ 0.13 equivalents per binding site) without addition of exogenous heme. Heme was removed prior to heme binding titrations using HPLC. Binding titrations using UV/vis spectroscopy revealed a stoichiometry of two hemes binding to each DH protein with a dissociation constant of $<10 \text{ nM}$ (Figure 6).

CD measurements show the DH maquette to be highly α -helical and thermostable. Its T_M is 71°C in the apo-state and greater than boiling when two hemes B are bound. Addition of heme B decreases the measured ellipticity at 222 nm at 25°C from $-27,900 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ in the apo-state to $-21,100 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ res}^{-1}$ when two hemes per protein are present, a 24% change (Figure 6C). This loss of ellipticity upon heme binding is similar to that observed in the RC maquette, but it stands in contrast to many other designed porphyrin-binding four-helix bundle proteins in which cofactor binding intensifies the ellipticity at 222 nm, presumably by promoting folding (Ghirlanda et al., 2004; Rau and Haehnel, 1998; Polizzi et al., 2017; Bender et al., 2007; Hutchins et al., 2020). One exception is the highly stable 4PA maquette, in which a small decrease in ellipticity at 222 nm upon porphyrin binding was attributed to a porphyrin transition (McAllister et al., 2008). Given that crystal structures of the RC maquette show comparable α -helix content with and without heme (*vide infra*), the decreased ellipticity at 222 nm upon heme binding may represent a porphyrin transition or a subtle rearrangement of the tertiary structure rather than loss of secondary structure.

The DH maquette exhibits many properties similar to the *P/A* module of the RC maquette. Both have high α -helix content, increase their melting temperatures to above boiling upon cofactor binding, have comparable heme-induced changes in their CD spectra, bind heme in *E. coli* and retain it through the purification process, and bind porphyrins stoichiometrically with low-nM dissociation constants or better. While the high-resolution structure of the DH maquette has not yet been solved experimentally, machine learning methods predict accurate folding. These results indicate that the *P/A* module of the RC maquette is capable of functioning as an independent protein that does not rely on the *D* module.

Crystal structures of RC maquette in various stages of cofactor assembly

The RC maquette was intended to fold as a canonical left-handed coiled coil with antiparallel α -helices. Since the

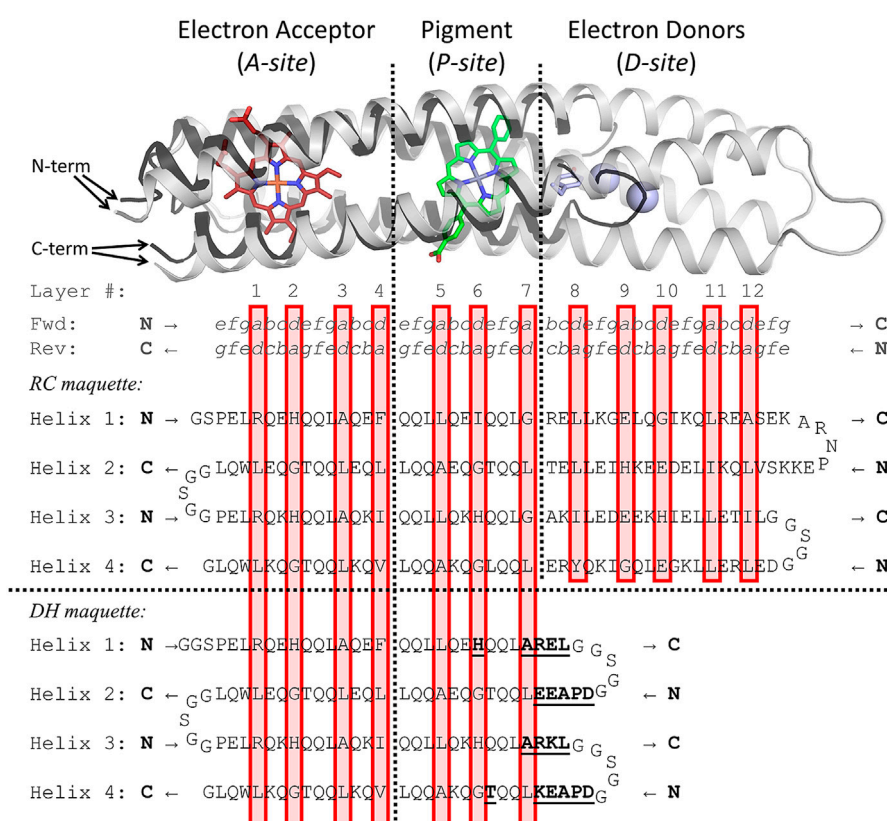


FIGURE 5

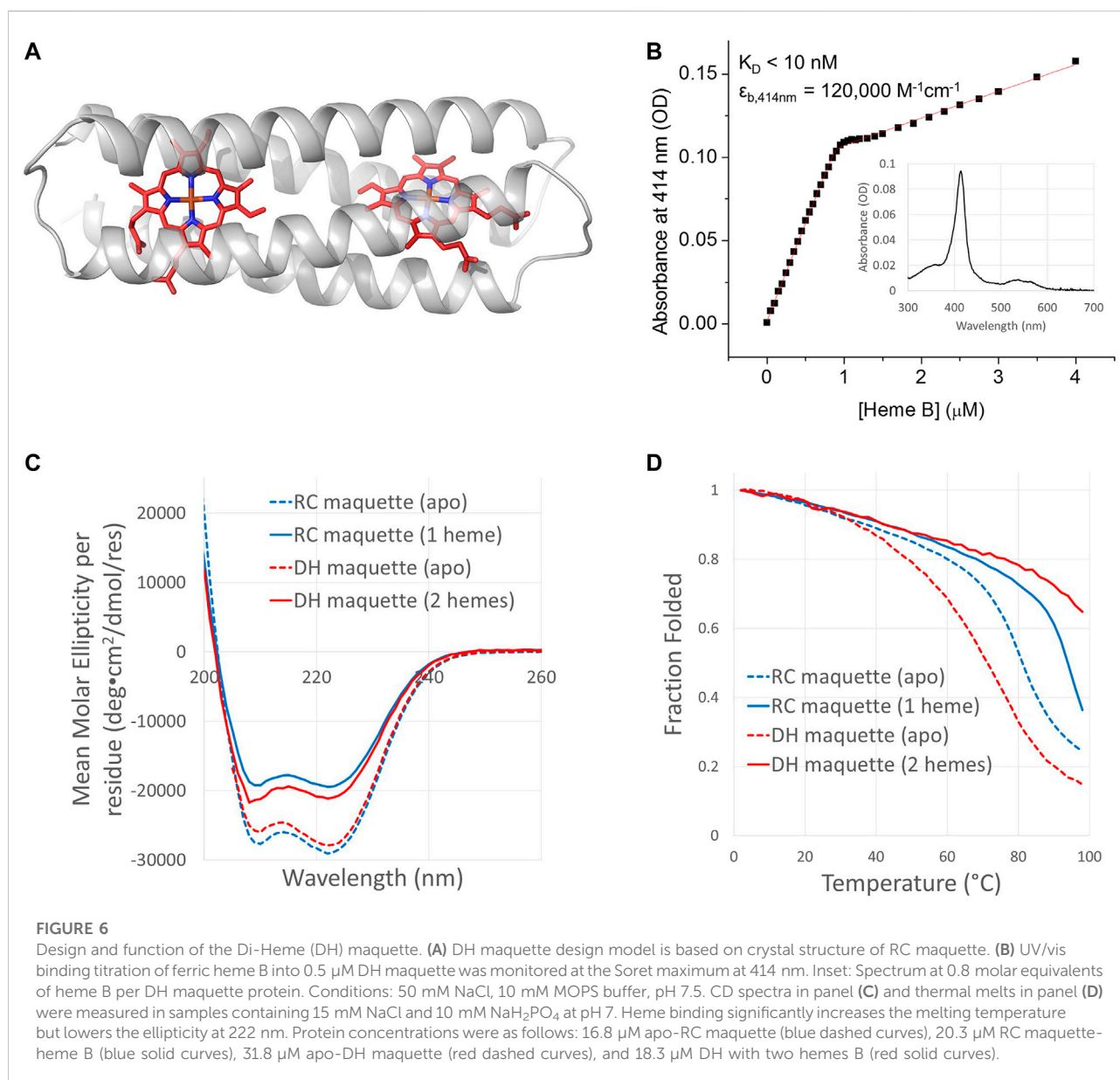
Sequence and structure comparison of RC and DH maquettes. Above, an RC maquette model based on crystal structures in white helices is overlaid with the RoseTTAFold (model 1) predicted structure of the DH maquette in black helices. The RoseTTAFold prediction has a high confidence score of 0.85. (AlphaFold predicted models are similar). The sequence diagrams below align the RC and DH maquette sequences. Differences between the maquettes in helical segments are indicated in bold and underlined in the DH maquette sequence. As in Figure 2, coiled coil heptad positions *a* through *g* are denoted in the top two rows, and *a/d* layers of the hydrophobic core are indicated by red boxes. The amino acid sequences have each helix written on a new line with helices 2 and 4 written in reverse from C-terminus to N-terminus as indicated by arrows and bold N or C at the ends of each helix. (Full forward sequences are given in Supplementary Table S1).

maquette was designed with minimal use of computational methods, we did not produce an atomistic design model *a priori* to which we can compare crystal structures. To evaluate the accuracy of our design methods, we fitted the RC maquette crystal structure to Crick parameters using the Coiled-Coil Crick Parameterization (CCCP) tool (Grigoryan and Degrado, 2011; Crick, 1953a). The amino acid backbone in the helical region of the holo-state RC maquette crystal structure (PDB ID: 5VJS) deviates from an idealized four-helix coiled coil by 0.96 Å C_α RMSD. The crystal structure averages 3.494 residues per α-helical turn, very close to the intended 3.5 residues per turn. Other crystal structures of the RC maquette have similar Crick parameters. The agreement between the overall fold and an idealized coiled coil reflects our adherence to well-established design rules and the natural

predisposition of helical bundles toward forming left-handed coiled coils.

Structure and flexibility of electron acceptor binding site

The *P/A* module of the RC maquette with heme B and ZnPCP bound (PDB ID: 5VJS) matches the cyt. *b₆* subunit of cyt. *b₆f* (PDB ID: 6RQF) to within 2.6 Å C_α RMSD, despite having only 14% sequence identity in the 125-residue aligned region (Malone et al., 2019) (Figure 7). Residues His9 and His110 in the RC maquette coordinate the Fe(III) of the heme ligand with their N_ε atoms; the His N_δ atoms are backed up by hydrogen bonds to second-shell Thr residues as designed. The notch Gly residues make the intended van der Waals contacts with the porphyrin rings, and *a/d* layers of the hydrophobic



core in the *P/A* module resemble those of cyt. *b* and *b₆* as intended.

RC maquette crystal structures reveal a highly accurate design, but they also provide evidence of conformational flexibility in the electron acceptor site. Heme B has a pseudo- C_2 symmetry axis passing through the α and γ meso positions, but the symmetry of the molecule is broken by substituent vinyl groups (Figure 8A). Similar to some native heme proteins, the RC maquette does not strongly discriminate between the two possible 180° rotations about the α - γ meso axis (Mortuza and Whitford, 1997; Vallone et al., 2004). Even at 1.45 Å resolution (PDB ID: 5VJT), the positions of the vinyl groups are ambiguous (see Figure 8B).

Further evidence of conformational flexibility in the heme site is found in an RC maquette crystal structure with a Leu71His mutation (PDB ID: 5VJU). Relative to Leu71 structures, the mutant exhibits a 90° rotation of the heme about an axis passing through the heme iron and His N_ϵ atoms (Figures 8B,C). Despite the rotation, protein-cofactor contacts as well as inter-cofactor distances are unchanged. The cause of the change in heme conformation is unclear. The Leu71His mutation may trigger a subtle conformational shift that is translated more than 2 nm from the *D* module to the *A*-site. However, the most obvious conformational difference in the *A*-site of the mutant is a change in crystal contacts involving the heme propionate groups, suggesting that the two heme binding orientations are similar

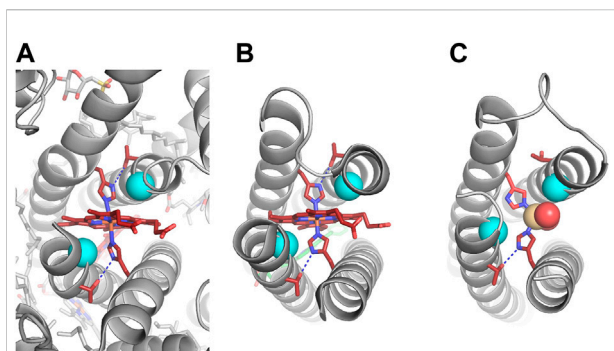


FIGURE 7

Experimental structures with and without heme. Heme B, ligating His residues, and second-shell Thr residues are shown in dark red. Notch Gly residues that make van der Waals contact with the porphyrin ring have C_{α} atoms shown in cyan. Hydrogen bonds are indicated by dashed blue lines. (A) Cryo-electron microscopy structure of spinach cytochrome b_6f (PDB ID: 6RQF) (Malone et al., 2019). (B) X-ray crystal structure of RC maquette with heme B and a Zn porphyrin (PDB ID: 5VJS) (Ennist et al., 2022). (C) X-ray crystal structure of RC maquette without heme B at 1.78 Å resolution (PDB ID: 8D9O). Cd(II) is shown as a tan sphere coordinated by the two histidines and two water molecules (red spheres). In the heme-free structure, the A-site is narrower, and the imidazole ring of His110 is flipped, breaking the hydrogen bond to Thr91.

in energy in solution but that different crystal packing arrangements favor one state over the other.

A complete understanding of cofactor binding requires consideration of the apo-state structure and its transition to the holo-state. While the RC maquette crystallized more readily with heme bound, one heme-free structure was solved (PDB ID: 8D9O), which showed that the maquette retains its α -helical structure in the absence of heme (Figure 7C). The heme-bound and heme-free structures align with a C_{α} RMSD of 1.2 Å when inter-helical loops are not included in the alignment. The heme-free structure was obtained using a crystallization condition that included a large excess of $CdCl_2$. Several Cd(II) ions are resolved in the structure, one of which effectively takes the place of the heme by coordinating the two A-site His residues as well as two water molecules. An additional water molecule in the A-site is hydrogen bonded to Gln111, and the side chains of Trp, Gln, and Leu residues adopt rotamers that would clash with the heme cofactor were it present. Measurement of the superhelical radius near the A-site using CCCP (Grigoryan and Degradó, 2011) indicates that the bundle widens to 7.8 Å radius in the heme-bound RC maquette (PDB ID: 5VJT) compared to 7.3 Å when Cd(II) occupies the A-site (PDB ID: 8D9O), a 7% change. Only one of the two His-Thr pairs in the A-site forms a hydrogen bond in the heme-free structure, whereas both hydrogen bonds form when heme is present. These structural differences suggest that the cofactor binding site is only partially pre-organized prior to heme binding.

Given that the a/d core of the RC maquette near the A-site includes two Gly, two His, and two Arg residues, all of which are likely to be strongly destabilizing in the apo-state, perhaps it is unsurprising that crystallization attempts without a stabilizing A-site ligand were unsuccessful. In a true apo-state where the A-site His residues do not coordinate a metal of any kind, it is likely that the A-site is more disordered. The decreased thermostability of the apo-protein (Figure 6) supports this conclusion.

While pre-organization of cofactor binding sites has been shown to aid binding specificity in some designed proteins (Ruckthong et al., 2016; Mills et al., 2013), pre-organization is not a universal requirement for high-affinity cofactor binding. Many native heme proteins employ a “partial fold option” for heme binding in which a partly-structured apo-protein with a disordered heme binding region collapses into a fully native-like structure only upon cofactor binding (Falzone et al., 1996; Landfried et al., 2007). This cofactor assembly strategy balances the apo-state stability needed to minimize the entropic cost of binding with the flexibility required for insertion of a bulky, hydrophobic porphyrin cofactor. Without the use of lyases to aid cofactor attachment as in the assembly of many biliproteins (Schluchter et al., 2010; Mancini et al., 2018), excessive stability might impede cofactor binding by requiring the disruption of strong protein-protein interactions to allow ligand entry into the binding pocket. Indeed, there is little pre-organization of heme sites in the apo-forms of myoglobin (Lin et al., 1994), cyt. b_5 (Falzone et al., 1996), cyt. b_{562} (Feng et al., 1994), and cyt. P450_{cam} (Pfeil et al., 1993), all of which transition from a partially folded apo-state to a more structured, native-like holo-state. In the RC maquette, the high heme affinity

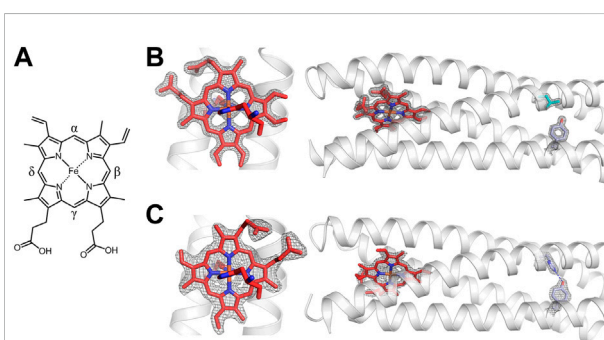
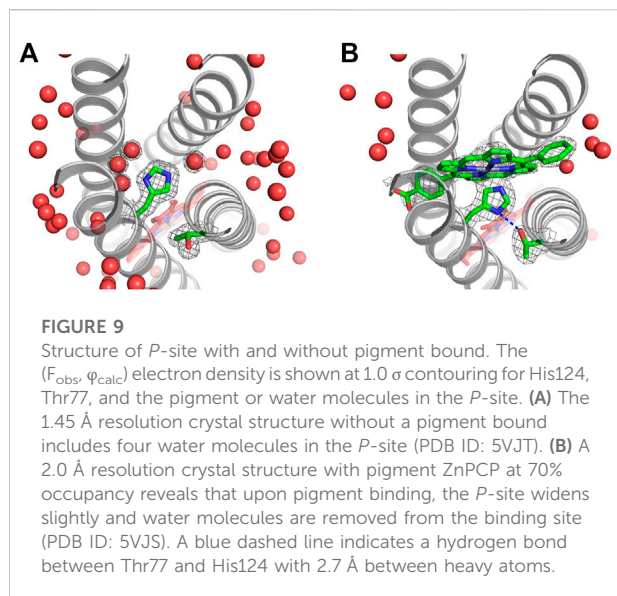


FIGURE 8

A Leu71His mutation distant from the A-site triggers a 90° rotation of heme B. (A) Chemical structure of heme B. (B) The 1.45 Å resolution crystal structure with heme B and Leu71 (PDB ID: 5VJT). Left: close-up of A-site is shown with 2.0 σ contouring of (F_{obs} , φ_{calc}) map around heme B. Right: larger context of same structure shows heme B, Leu71, and Tyr168 with 2.0 σ contouring. (C) The 2.08 Å resolution crystal structure with heme B and a His71-Tyr168 hydrogen bond (PDB ID: 5VJU). Left: close-up of A-site in Leu71His mutant is shown with 1.0 σ contouring of (F_{obs} , φ_{calc}) map around heme B. Right: larger context of same structure shows heme B and His71-Tyr168 hydrogen bond (indicated by blue dashed line) at 1.5 σ contouring.



and mostly folded but flexible apo-state indicate that the partial fold option has been applied successfully in a *de novo*-designed heme-binding protein.

Structure of pigment binding site

The RC maquette *A*-site lends itself to the partial fold option as a strategy for cofactor binding because it is close to the end of the bundle where the helices tend to be more dynamic and can more easily widen to accept a bulky cofactor. In contrast, the *P*-site is located in the more rigid center of the protein. In the 2.0 Å resolution RC maquette crystal structure with ZnPCP (PDB ID: 5VJS), omit maps confirm the presence of the pigment in the *P*-site, and Multi-wavelength Anomalous Dispersion (MAD) phasing positively identifies the Zn(II) ion in the ZnPCP cofactor (Ennist et al., 2022). Four *P*-site Gly residues are meant to create space for the Zn porphyrin pigment, increase backbone flexibility, and prevent a premature hydrophobic collapse of the binding cleft that could obstruct cofactor binding. Without a Zn porphyrin pigment (PDB IDs: 5VJT, 5VJU, 8D9O, and 8D9P), four structured water molecules occupy the *P*-site; as intended, the binding cavity does not collapse in the apo-state (see Figure 9). In all RC maquette crystal structures, temperature factors of the *P*-site Gly residues are comparable to other residues in the vicinity of the site, suggesting that the Gly are not significantly mobile. However, upon binding of ZnPCP (PDB ID: 5VJS), the superhelical radius within one α -helical turn of the *P*-site expands by 4% (0.3 Å) as measured by CCCP (Grigoryan and Degrado, 2011), demonstrating moderate flexibility in the helical backbone.

The electron density for the ZnPCP pigment is not as well-defined as that for the heme B in the same crystal structure.

This may be due to sub-stoichiometric occupancy of the ZnPCP (the pigment is assigned a 70% occupancy in the deposited structure). Alternatively, the ZnPCP may be more mobile than the heme because of the shape of the binding pocket or the fact that it has only one axial His ligand (His124) instead of two. Most natural proteins that use His imidazole side chains to bind metallotetrapyrroles such as chlorophylls and hemes use the His N_{ϵ} atom to coordinate the metal, probably because of steric limitations. However, there are rare examples of His N_{δ} ligation of hemes and chlorophylls in nature (Iverson et al., 1998; Umena et al., 2011; Jordan et al., 2001) and in earlier maquettes (Choma et al., 1994; Zhuang et al., 2004). Heme binding in the RC maquette *A*-site employs His N_{ϵ} atoms for heme iron ligation as designed (Figure 7). In the *P*-site, the electron density for the His124 ligand to ZnPCP suggests either a strained conformation or a mixture of states, such as ZnPCP-bound and unbound (PDB ID: 5VJS). The crystal structure is modeled with N_{ϵ} coordination of the ZnPCP in agreement with the design, but the geometry of ligation is not ideal. N_{δ} ligation is a poorer fit to the electron density, but remains a possibility. As modeled in the crystal structure, His124 positions its N_{δ} atom 2.7 Å from the hydroxyl oxygen of Thr77, suggesting a strong hydrogen bond when ZnPCP is bound. In contrast, crystal structures without ZnPCP clearly show that Thr77 is in a different rotameric state, and instead forms a hydrogen bond with the backbone carbonyl group of Thr73.

The process of Zn porphyrin pigment binding to the RC maquette *P*-site includes removal of four water molecules from the binding cavity, a 4% expansion of the superhelical radius around the *P*-site, and likely formation of a His-Thr hydrogen bond. The high ZnPCP affinity suggests that these modest conformational changes do not significantly impede pigment binding.

Structure of electron donor site

Crystal structures of the RC maquette show excellent agreement with the structure of bacterioferritin. The RC maquette *D* module with Cd(II) differs from Cd(II)-bacterioferritin by 1.25 Å C_{α} RMSD over a 96-residue aligned region (PDB IDs: 8D9O and 4CVS, respectively) (Hingorani et al., 2014). Both Cd(II) ions in the *D*-site have octahedral geometry. His67 and His138 coordinate the metal ions with their N_{δ} atoms and use their N_{ϵ} atoms to make hydrogen bonds with second-shell Asp residues as designed. Two axial bidentate Glu residues (Glu34 and Glu161) and two bridging Glu (Glu64 and Glu135) ligate the Cd(II) ions as shown in Supplementary Figure S2D. A bridging water or hydroxide ion coordinates both Cd(II) ions. The Tyr168 hydroxyl group makes a hydrogen bond with Glu34, and in the Leu71His mutant (PDB ID: 5VJU), Tyr168 also hydrogen bonds His71 as designed (Ennist et al., 2022).

A 1.90 Å-resolution Mn(II)-containing structure was obtained with heme B as the electron acceptor and no pigment (PDB ID: 8D9P). In this structure, the coordination geometry for the two Mn(II) ions matches that of the Cd(II)-containing structures, but with 5.0 Å between the Mn(II) ions. In contrast, the Cd(II) structures with and without heme B have the metals 3.6 Å apart (Supplementary Figures S2C,D). Despite this difference, crystallographic and spectroscopic data for RC maquette binding of Cd(II), Mn(II), and Co(II) support the conclusion that five amino acid ligands coordinate each metal ion.

While the Cd(II) and Mn(II) structures (grown at pH 4.6–5.9) are consistent with the design, crystal structures with Zn(II) show a different coordination geometry. All RC maquette crystals grew under acidic conditions. Given that protons are known to compete with metal binding in Due Ferri proteins (Pasternak et al., 2001), it is possible that acid affects the *D*-site in some structures. Crystals grown with Zn(II) at pH 4.4–4.5 (PDB IDs: 5VJS and 5VJT) have only one tetrahedral Zn(II) ion in the intended di-metal site (Ennist et al., 2022). The metal site furthest from the *P*-site is vacant, but a second tetrahedral Zn(II) ion is present on the outer surface of the RC maquette coordinated by the His138 N_ε atom and the two Asp residues that were intended as second-shell ligands.

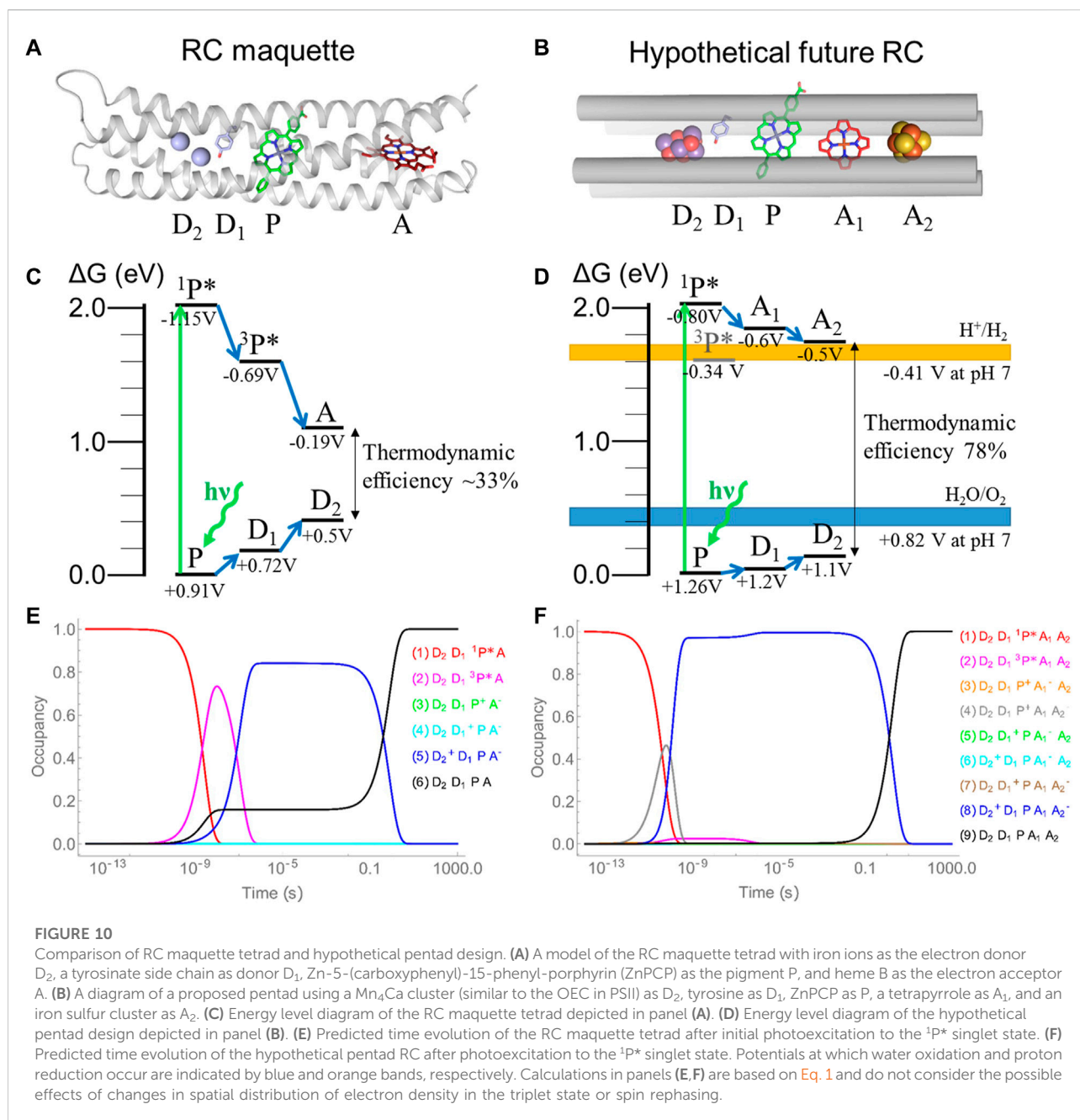
Comparison of DF3, which has four active site Gly residues, with other Due Ferri maquettes that lack helical Gly residues showed that a flexible binding pocket can enhance affinity for various metals (Torres Martin de Rosales et al., 2010). However, some studies have argued that high stability and pre-organization in a designed metal site can improve selectivity to avoid off-target binding geometries (Ruckthong et al., 2016; Mills et al., 2013). In the RC maquette, metals bind with high affinity, but the four Gly residues in the *D*-site may be partly responsible for the observed conformational heterogeneity in different crystal structures. This suggests a tradeoff wherein a highly pre-organized binding site may strictly preserve the intended structure at the cost of binding affinity, and a flexible metal binding site may have high affinity but risk distortion of the structure.

Critique of RC maquette and strategy to improve performance

The achievement of long-lived charge separation in the RC maquette represents an important step toward the *de novo* design of proteins for solar-to-fuel energy conversion (Ennist et al., 2022). However, several significant modifications need to be made before designed RCs become useful for renewable energy production. Future designs must avoid the use of triplet excited states, increase the thermodynamic efficiency and quantum yield of charge separation, and catalyze redox reactions for fuel production.

The RC maquette exploits a Zn tetrapyrrole pigment with a relatively high-yield photoexcited triplet state. Promotion from the ground state to the singlet excited state is followed by intersystem crossing to the triplet excited state on a nanosecond timescale. The triplet state is a useful tool in the development of the RC maquette, because its millisecond-scale lifetime enables slow electron transfer over a long distance, allowing for easier characterization (Figures 10A,C,E). However, triplet states present many challenges, including susceptibility to oxidative damage by singlet oxygen and energy losses due to intersystem crossing. It has often been noted that electron density can be more spatially confined in triplet states than in singlet states (Richert et al., 2017), which will tend to slow initial charge separation. In addition, once the electron is transferred to heme, spin rephasing to a singlet charge separated state may lead to faster charge recombination. Indeed, transient absorption measurements of the RC maquette (Ennist et al., 2022) show that charge separation takes place even slower and with lower yield than the calculated electron-transfer behavior in Figure 10E, which does not take into account the shape of the triplet state electron cloud or spin rephasing effects. Additional inefficiencies in the RC maquette result from large potential differences between redox centers; the RC maquette tetrad charge separated state has a free energy difference of ~0.69 eV between the oxidized Fe(III) donor and reduced ferrous heme B acceptor (Figure 10C), a ~33% thermodynamic yield from the excited singlet state energy of 2.06 eV. By comparison, the thermodynamic efficiency of PSII is ~44–56% (Vinyard et al., 2013).

Electron-tunneling calculations using Eq. 1 suggest that a much higher efficiency is attainable in a designed pentad. With ideal placement of electron-transfer cofactors and finely tuned redox potentials, the proposed pentad design in Figures 10B,D,F is predicted to have a 78% thermodynamic efficiency, a 99.5% quantum yield at 1 ms after excitation, and a charge separated state lifetime of 1.9 s. Its midpoint potentials for the tyrosine donor D₁, metal cluster donor D₂, and the pigment P/P^{•+} redox couple, are set to the same potentials as those in PSII (Vinyard et al., 2013). Lifetimes of intersystem crossing, internal conversion, and fluorescence are approximated as values typical of Zn porphyrins (Magdaong et al., 2020). The hypothetical pentad RC avoids the pigment triplet excited state by using a relatively short P-to-A₁ distance of 5.7 Å to facilitate rapid electron transfer from the singlet state that outcompetes intersystem crossing. A secondary electron acceptor can prolong the lifetime of charge separation (Moser et al., 2017). In Figure 10B,D,F, we conceive A₁ as a tetrapyrrole with a midpoint potential of –0.6 V and A₂ as an iron-sulfur cluster with a midpoint of –0.5 V vs. SHE. The high thermodynamic yield gives the charge separated state enough energy for both water oxidation and hydrogen evolution in the same RC. A summary of parameters



used for calculating electron-tunneling rates and efficiencies is given in [Supplementary Table S5](#).

Certainly there are many challenges in the development of a pentad RC such as that proposed in [Figure 10](#). Foremost among them is the photoassembly of an oxygen evolving cluster (OEC). The Mn_4CaO_5 OEC of PSII is assembled by successive cycles of metal ion binding and photooxidation ([Bao and Burnap, 2016](#)). In higher oxidation states, metals are more likely to form metal oxide clusters, so it may be possible to use an RC maquette to emulate the process of OEC photoassembly. Other major hurdles include tuning midpoint potentials, engineering catalytic sites, insulating the reducing and

oxidizing sites to prevent inter-protein charge recombination or generation of reactive oxygen species, and *in vivo* self-assembly. A solution to these many challenges will require iterative cycles of design, characterization, and analysis to bring us closer to goal of enhanced solar energy conversion through protein design.

Discussion

We have demonstrated a method of designing helical bundle proteins based on human intuition and experimental evidence of

design rules from native and designed proteins. Our approach is sufficient to achieve high-level functions including multi-cofactor binding and photochemical charge separation. New machine learning methods including AlphaFold (Jumper et al., 2021) and RoseTTAFold (Baek et al., 2021) are likely to increase the robustness of this design process, but a firm intuitive foundation will facilitate the effective use of these novel computational tools. We have shown that high-precision rational design of complex four-helix bundles is feasible for human designers without the aid of computational methods.

The RC maquette is a useful testbed to explore photochemical charge separation. Limited promiscuity in cofactor binding allows assembly with different electron donors, pigments, and acceptors that span a range of redox potentials. Crystal structures show that the hydrogen bonding environment involving a redox-active tyrosine can be modified to examine PSII-like tyrosine oxidation. The modular design is amenable to the future reengineering of electron-transfer distances. These attributes enable the same protein scaffold to be used for investigation of electron-transport chains that have many different properties.

The successful modular design and characterization of an RC maquette sets the stage for future functional improvements. We have shown that in principle, a designed RC with a cofactor pentad could trap a charge-separated state with enough energy for water splitting in a single photosystem and with a thermodynamic efficiency surpassing that of natural photosystems. A new RC design with binding sites for five redox centers will need to balance flexibility in cofactor binding pockets with enough apo-state stability to minimize the entropic cost of transitioning to the holo-state. Continued use of the maquette approach to iteratively improve RC maquette thermodynamic and quantum efficiencies, fine tune redox potentials, and introduce catalytic activity may one day lead to new photosynthetic pathways for efficient solar fuel production.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

NE, LD, and CM conceived the project. NE designed the RC maquette and DH maquette proteins. NE performed cofactor binding titrations and other biochemical assays. NE grew protein

crystals. SS and NE collected and processed X-ray diffraction data and solved crystal structures. NE performed electron-transfer rate calculations with contributions from CM. NE wrote the manuscript with contributions from CM, SS, and LD.

Funding

Funding was provided by the Department of Energy, Office of Science, Office of Basic Energy Sciences, through the Photosynthetic Antenna Research Center (PARC), an Energy Frontier Research Center, under Award DESC0001035. Additional funding was provided by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0001543.

Acknowledgments

We thank Tatiana V. Espipova, the laboratory of Sergei A. Vinogradov, and the laboratory of Jonathan S. Lindsey for generously providing synthetic tetrapyrrole pigments used in this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.997295/full#supplementary-material>

References

- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). Phenix: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* 66 (2), 213–221. doi:10.1107/S0907444909052925
- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., et al. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* 68 (4), 352–367. doi:10.1107/S0907444912001308
- Alcala-Torano, R., Halloran, N., Gwerder, N., Sommer, D. J., and Ghirlanda, G. (2021). Light-driven CO₂ reduction by Co-cytochrome b 562. *Front. Mol. Biosci.* 8, 609654. doi:10.3389/fmolb.2021.609654
- Aravindu, K., Mass, O., Vairaprakash, P., Springer, J. W., Yang, E., Niedzwiedzki, D. M., et al. (2013). Amphiphilic chlorins and bacteriochlorins in micellar environments. Molecular design, de novo synthesis, and photophysical properties. *Chem. Sci.* 4, 3459–3477. doi:10.1039/c3sc51335a
- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373 (6557), 871–876. doi:10.1126/science.abj8754
- Bao, H., and Burnap, R. L. (2016). Photoactivation: The light-driven assembly of the water oxidation complex of photosystem II. *Front. Plant Sci.* 7, 578. doi:10.3389/fpls.2016.00578
- Barber, J. (2016). The Mn₄Ca-cluster of the photosynthetic oxygen evolving centre: Its structure, function and evolution. *Biochemistry* 55, 1. doi:10.1021/acs.biochem.6b00794
- Barber, J., and Tran, P. D. (2013). From natural to artificial photosynthesis. *J. R. Soc. Interface* 10 (81), 20120984. doi:10.1098/rsif.2012.0984
- Bender, G. M., Lehmann, A., Zou, H., Cheng, H., Fry, H. C., Engel, D., et al. (2007). De novo design of a single-chain diphenylporphyrin metalloprotein. *J. Am. Chem. Soc.* 129 (35), 10732–10740. doi:10.1021/ja071199j
- Berry, E. A., and Trumpower, B. L. (1987). Simultaneous determination of hemes a, b, and c from pyridine heme spectra. *Anal. Biochem.* 161 (1), 1–15. doi:10.1016/0003-2697(87)90643-9
- Bertini, I., and Luchinat, C. (1984). High spin cobalt(II) as a probe for the investigation of metalloproteins. *Adv. Inorg. Biochem.* 6, 71–111.
- Betz, S. F., and DeGrado, W. F. (1996). Controlling topology and native-like behavior of de novo-designed peptides: Design and characterization of antiparallel four-stranded coiled coils. *Biochemistry* 35 (21), 6955–6962. doi:10.1021/bi960095a
- Betz, S. F., Liebman, P. A., and DeGrado, W. F. (1997). De novo design of native proteins: Characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles. *Biochemistry* 36 (9), 2450–2458. doi:10.1021/bi961704h
- Blankenship, R. E., Tiede, D. M., Barber, J., Brudvig, G. W., Fleming, G., Ghirardi, M., et al. (2011). Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* 332 (6031), 805–809. doi:10.1126/science.1200165
- Boyken, S. E., Chen, Z., Groves, B., Langan, R. A., Oberdorfer, G., Ford, A., et al. (2016). De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* 352 (6286), 680–687. doi:10.1126/science.aad8865
- Cardona, T. (2016). Reconstructing the origin of oxygenic photosynthesis: Do assembly and photoactivation recapitulate evolution? *Front. Plant Sci.* 7, 257. doi:10.3389/fpls.2016.00257
- Chino, M., Maglio, O., Nastri, F., Pavone, V., DeGrado, W. F., and Lombardi, A. (2015). Artificial diiron enzymes with a de novo designed four-helix bundle structure. *Eur. J. Inorg. Chem.* 2015 (21), 3371–3390. doi:10.1002/ejic.201500470
- Choma, C. T., Lear, J. D., Nelson, M. J., Dutton, P. L., Robertson, D. E., and DeGrado, W. F. (1994). Design of a heme-binding four-helix bundle. *J. Am. Chem. Soc.* 116 (3), 856–865. doi:10.1021/ja00082a005
- Cochran, F. V., Wu, S. P., Wang, W., Nanda, V., Saven, J. G., Therien, M. J., et al. (2005). Computational de novo design and characterization of a four-helix bundle protein that selectively binds a nonbiological cofactor. *J. Am. Chem. Soc.* 127 (5), 1346–1347. doi:10.1021/ja044129a
- Conlan, B., Cox, N., Su, J. H., Hillier, W., Messinger, J., Lubitz, W., et al. (2009). Photo-catalytic oxidation of a di-nuclear manganese centre in an engineered bacterioferritin 'reaction centre'. *Biochim. Biophys. Acta* 1787 (9), 1112–1121. doi:10.1016/j.bbabi.2009.04.011
- Crick, F. H. C. (1953a). The Fourier transform of a coiled-coil. *Acta Crystallogr.* 6, 685–689. doi:10.1107/s0365110x53001952
- Crick, F. H. C. (1953b). The packing of alpha-helices - simple coiled-coils. *Acta Crystallogr.* 6 (8–9), 689–697. doi:10.1107/s0365110x53001964
- Dautant, A., Meyer, J. B., Yariv, J., Precigoux, G., Sweet, R. M., Kalb, A. J., et al. (1998). Structure of a monoclinic crystal from of cytochrome b1 (Bacterioferritin) from *E. coli*. *Acta Crystallogr. D. Biol. Crystallogr.* 54 (1), 16–24. doi:10.1107/s0907444997006811
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* 66 (4), 486–501. doi:10.1107/S0907444910007493
- Ennist, N. M., Mancini, J. A., Auman, D. B., Bialas, C., Iwanicki, M. J., Esipova, T. V., et al. (2017). "Maquette strategy for creation of light- and redox-active proteins," in *Photosynthesis and bioenergetics*. Editors J. Barber and A. V. Ruban (World Scientific Publishers), 1–33.
- Ennist, N. M., Zhao, Z., Stayrook, S. E., Discher, B. M., Dutton, P. L., and Moser, C. C. (2022). De novo protein design of photochemical reaction centers. *Nat. Commun.* 13 (1), 4937. doi:10.1038/s41467-022-32710-5
- Esposito, M. D., De Vries, S., Crimi, M., Ghelli, A., Patarnello, T., and Meyer, A. (1993). Mitochondrial cytochrome b: Evolution and structure of the protein. *Biochim. Biophys. Acta* 1143 (3), 243–271. doi:10.1016/0005-2728(93)90197-n
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D. Biol. Crystallogr.* 62 (1), 72–82. doi:10.1107/S0907444905036693
- Faiella, M., Andreozzi, C., de Rosales, R. T., Pavone, V., Maglio, O., Nastri, F., et al. (2009). An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.* 5 (12), 882–884. doi:10.1038/nchembio.257
- Falzone, C. J., Mayer, M. R., Whiteman, E. L., Moore, C. D., and Lecomte, J. T. (1996). Design challenges for hemoproteins: The solution structure of apocytochrome b5. *Biochemistry* 35 (21), 6519–6526. doi:10.1021/bi960501q
- Farid, T. A., Kodali, G., Solomon, L. A., Lichtenstein, B. R., Sheehan, M. M., Fry, H. C., et al. (2013). Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* 9 (12), 826–833. doi:10.1038/nchembio.1362
- Feng, Y., Sligar, S. G., and Wand, A. J. (1994). Solution structure of apocytochrome b562. *Nat. Struct. Biol.* 1 (1), 30–35. doi:10.1038/nsb0194-30
- Fry, H. C., Lehmann, A., Saven, J. G., DeGrado, W. F., and Therien, M. J. (2010). Computational design and elaboration of a de novo heterotetrameric alpha-helical protein that selectively binds an emissive abiological (porphyrinato)zinc chromophore. *J. Am. Chem. Soc.* 132 (11), 3997–4005. doi:10.1021/ja907407m
- Gernert, K. M., Surles, M. C., Labean, T. H., Richardson, J. S., and Richardson, D. C. (1995). The alacoli: A very tight, antiparallel coiled-coil of helices. *Protein Sci.* 4 (11), 2252–2260. doi:10.1002/pro.5560041102
- Ghirlanda, G., Osyczka, A., Liu, W., Antolovich, M., Smith, K. M., Dutton, P. L., et al. (2004). De novo design of a D₂-symmetrical protein that reproduces the diheme four-helix bundle in cytochrome bc₁. *J. Am. Chem. Soc.* 126 (26), 8141–8147. doi:10.1021/ja039935g
- Gibney, B. R., Rabanal, F., Skalicky, J. J., Wand, A. J., and Dutton, P. L. (1999). Iterative protein redesign. *J. Am. Chem. Soc.* 121 (21), 4952–4960. doi:10.1021/ja9833117
- Glover, S. D., Jorge, C., Liang, L., Valentine, K. G., Hammarstrom, L., and Tommos, C. (2014). Photochemical tyrosine oxidation in the structurally well-defined α3Y protein: Proton-coupled electron transfer and a long-lived tyrosine radical. *J. Am. Chem. Soc.* 136 (40), 14039–14051. doi:10.1021/ja503348d
- Grayson, K. J., and Anderson, J. L. R. (2018). The ascent of man (made oxidoreductases). *Curr. Opin. Struct. Biol.* 51, 149–155. doi:10.1016/j.sbi.2018.04.008
- Grigoryan, G., and DeGrado, W. F. (2011). Probing designability via a generalized model of helical bundle geometry. *J. Mol. Biol.* 405 (4), 1079–1100. doi:10.1016/j.jmb.2010.08.058
- Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993). A switch between two- and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262 (5138), 1401–1407. doi:10.1126/science.8248779
- Hingorani, K., Pace, R., Whitney, S., Murray, J. W., Smith, P., Cheah, M. H., et al. (2014). Photo-oxidation of tyrosine in a bio-engineered bacterioferritin 'reaction centre'-a protein model for artificial photosynthesis. *Biochim. Biophys. Acta* 1837 (10), 1821–1834. doi:10.1016/j.bbabi.2014.07.019
- Huang, P. S., Oberdorfer, G., Xu, C., Pei, X. Y., Nannenga, B. L., Rogers, J. M., et al. (2014). High thermodynamic stability of parametrically designed helical bundles. *Science* 346 (6208), 481–485. doi:10.1126/science.1257481

- Hunte, C., Koepke, J., Lange, C., Rossmanith, T., and Michel, H. (2000). Structure at 2.3 Å resolution of the cytochrome bc(1) complex from the yeast *Saccharomyces cerevisiae* co-crystallized with an antibody Fv fragment. *Structure* 8 (6), 669–684. doi:10.1016/s0969-2126(00)00152-0
- Hutchins, G. H., Noble, C. E., Blackburn, H., Hardy, B., Landau, C., Parnell, A. E., et al. (2020). Precision design of single and multi-heme de novo proteins. *bioRxiv*. doi:10.1101/2020.09.24.311514
- Imahori, H., Tamaki, K., Guldi, D. M., Luo, C., Fujitsuka, M., Ito, O., et al. (2001). Modulating charge separation and charge recombination dynamics in porphyrin-fullerene linked dyads and triads: Marcus-normal versus inverted region. *J. Am. Chem. Soc.* 123 (11), 2607–2617. doi:10.1021/ja003346i
- Iverson, T. M., Arciero, D. M., Hsu, B. T., Logan, M. S., Hooper, A. B., and Rees, D. C. (1998). Heme packing motifs revealed by the crystal structure of the tetra-heme cytochrome c554 from *Nitrosomonas europaea*. *Nat. Struct. Biol.* 5 (11), 1005–1012. doi:10.1038/2975
- Joosten, R. P., Long, F., Murshudov, G. N., and Perrakis, A. (2014). The PDB_REDO server for macromolecular structure model optimization. *IUCr* 1 (4), 213–220. doi:10.1107/S2052252514009324
- Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. (2001). Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* 411 (6840), 909–917. doi:10.1038/35082000
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596 (7873), 583–589. doi:10.1038/s41586-021-03819-2
- Kalman, L., LoBrutto, R., Allen, J. P., and Williams, J. C. (1999). Modified reaction centres oxidize tyrosine in reactions that mirror photosystem II. *Nature* 402 (6762), 696–699. doi:10.1038/45300
- Kamtekar, S., and Hecht, M. H. (1995). Protein motifs. 7. The four-helix bundle: What determines a fold? *Faseb J.* 9 (11), 1013–1022. doi:10.1096/fasebj.9.11.7649401
- Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M., and Hecht, M. H. (1993). Protein design by binary patterning of polar and nonpolar amino acids. *Science* 262 (5140), 1680–1685. doi:10.1126/science.8259512
- Kodali, G., Mancini, J. A., Solomon, L. A., Episova, T. V., Roach, N., Hobbs, C. J., et al. (2017). Design and engineering of water-soluble light-harvesting protein maquettes. *Chem. Sci.* 8 (1), 316–324. doi:10.1039/C6SC02417C
- Kumar, P., and Woolfson, D. N. (2021). Socket2: A program for locating, visualising, and analysing coiled-coil interfaces in protein structures. *Bioinformatics* 37, 4575–4577. doi:10.1093/bioinformatics/btab631
- Lahr, S. J., Engel, D. E., Stayrook, S. E., Maglio, O., North, B., Geremia, S., et al. (2005). Analysis and design of turns in alpha-helical hairpins. *J. Mol. Biol.* 346 (5), 1441–1454. doi:10.1016/j.jmb.2004.12.016
- Landfried, D. A., Vuletic, D. A., Pond, M. P., and Lecomte, J. T. (2007). Structural and thermodynamic consequences of b heme binding for monomeric apoglobins and other apoproteins. *Gene* 398 (1–2), 12–28. doi:10.1016/j.gene.2007.02.046
- Leman, J. K., Weitzner, B. D., Lewis, S. M., Adolf-Bryfogle, J., Alam, N., Alford, R. F., et al. (2020). Macromolecular modeling and design in Rosetta: Recent methods and frameworks. *Nat. Methods* 17 (7), 665–680. doi:10.1038/s41592-020-0848-2
- Lichtenstein, B. R., Farid, T. A., Kodali, G., Solomon, L. A., Anderson, J. L., Sheehan, M. M., et al. (2012). Engineering oxidoreductases: Engineering oxidoreductases: Maquette proteins designed from scratch. *Biochem. Soc. Trans.* 40 (3), 561–566. doi:10.1042/BST20120067
- Lieschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *Acta Crystallogr. D. Struct. Biol.* 75 (10), 861–877. doi:10.1107/S2059798319011471
- Lieschner, D., Afonine, P. V., Moriarty, N. W., Poon, B. K., Sobolev, O. V., Terwilliger, T. C., et al. (2017). Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D. Struct. Biol.* 73 (2), 148–157. doi:10.1107/S2059798316018210
- Lin, L., Pinker, R. J., Forde, K., Rose, G. D., and Kallenbach, N. R. (1994). Molten globular characteristics of the native state of apomyoglobin. *Nat. Struct. Biol.* 1 (7), 447–452. doi:10.1038/nsb0794-447
- Liu, J., Zheng, Q., Deng, Y., Li, Q., Kallenbach, N. R., and Lu, M. (2007). Conformational specificity of the lac repressor coiled-coil tetramerization domain. *Biochemistry* 46 (51), 14951–14959. doi:10.1021/bi701930d
- Lombardi, A., Pirro, F., Maglio, O., Chino, M., and DeGrado, W. F. (2019). De novo design of four-helix bundle metalloproteins: One scaffold, diverse reactivities. *Acc. Chem. Res.* 52 (5), 1148–1159. doi:10.1021/acs.accounts.8b00674
- Magdaong, N. C. M., Taniguchi, M., Diers, J. R., Niedzwiedzki, D. M., Kirmaier, C., Lindsey, J. S., et al. (2020). Photophysical properties and electronic structure of zinc(II) porphyrins bearing 0–4 meso-phenyl substituents: Zinc porphine to zinc Tetraphenylporphyrin (ZnTPP). *J. Phys. Chem. A* 124 (38), 7776–7794. doi:10.1021/acs.jpca.0c06841
- Malone, L. A., Qian, P., Mayneord, G. E., Hitchcock, A., Farmer, D. A., Thompson, R. F., et al. (2019). Cryo-EM structure of the spinach cytochrome b6 f complex at 3.6 Å resolution. *Nature* 575 (7783), 535–539. doi:10.1038/s41586-019-1746-6
- Mancini, J. A., Sheehan, M., Kodali, G., Chow, B. Y., Bryant, D. A., Dutton, P. L., et al. (2018). De novo synthetic biliprotein design, assembly and excitation energy transfer. *J. R. Soc. Interface* 15 (141), 20180021. doi:10.1098/rsif.2018.0021
- Mann, S. I., Nayak, A., Gassner, G. T., Therien, M. J., and DeGrado, W. F. (2021). De novo design, solution characterization, and crystallographic structure of an abiological Mn-Porphyrin-Binding protein capable of stabilizing a Mn(V) species. *J. Am. Chem. Soc.* 143 (1), 252–259. doi:10.1021/jacs.0c10136
- Marsh, E. N., and DeGrado, W. F. (2002). Noncovalent self-assembly of a heterotetrameric diiron protein. *Proc. Natl. Acad. Sci. U. S. A.* 99 (8), 5150–5154. doi:10.1073/pnas.052023199
- Mathews, F. S., Bethge, P. H., and Czerwinski, E. W. (1979). The structure of cytochrome b562 from *Escherichia coli* at 2.5 Å resolution. *J. Biol. Chem.* 254 (5), 1699–1706. doi:10.1016/s0021-9258(17)37829-8
- McAllister, K. A., Zou, H., Cochran, F. V., Bender, G. M., Senes, A., Fry, H. C., et al. (2008). Using alpha-helical coiled-coils to design nanostructured metalloporphyrin arrays. *J. Am. Chem. Soc.* 130 (36), 11921–11927. doi:10.1021/ja800697g
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40 (4), 658–674. doi:10.1107/S0021889807021206
- Megiatto, J. D., Jr., Mendez-Hernandez, D. D., Tejada-Ferrari, M. E., Teillout, A. L., Llansola-Portoles, M. J., Kodis, G., et al. (2014). A bioinspired redox relay that mimics radical interactions of the Tyr-His pairs of photosystem II. *Nat. Chem.* 6 (5), 423–428. doi:10.1038/nchem.1862
- Mersch, D., Lee, C. Y., Zhang, J. Z., Brinkert, K., Fontecilla-Camps, J. C., Rutherford, A. W., et al. (2015). Wiring of photosystem II to hydrogenase for photoelectrochemical water splitting. *J. Am. Chem. Soc.* 137 (26), 8541–8549. doi:10.1021/jacs.5b03737
- Mills, J. H., Khare, S. D., Bolduc, J. M., Forouhar, F., Mulligan, V. K., Lew, S., et al. (2013). Computational design of an unnatural amino acid dependent metalloprotein with atomic level accuracy. *J. Am. Chem. Soc.* 135 (36), 13393–13399. doi:10.1021/ja403503m
- Mirdita, M., Schütze, K., Moriawaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nat. Methods* 19 (6), 679–682. doi:10.1038/s41592-022-01488-1
- Mortuza, G. B., and Whitford, D. (1997). Mutagenesis of residues 27 and 78 modulates heme orientation in cytochrome b5. *FEBS Lett.* 412 (3), 610–614. doi:10.1016/s0014-5793(97)00870-3
- Moser, C. C., Anderson, J. L., and Dutton, P. L. (2010). Guidelines for tunneling in enzymes. *Biochim. Biophys. Acta* 1797 (9), 1573–1586. doi:10.1016/j.bbabi.2010.04.441
- Moser, C. C., Keske, J. M., Warncke, K., Farid, R. S., and Dutton, P. L. (1992). Nature of biological electron transfer. *Nature* 355 (6363), 796–802. doi:10.1038/355796a0
- Moser, C. C., Sheehan, M. M., Ennist, N. M., Kodali, G., Bialas, C., Englander, M. T., et al. (2016). De novo construction of redox active proteins. *Methods Enzymol.* 580, 365–388. doi:10.1016/bs.mie.2016.05.048
- Moser, C. C., Ennist, N. M., Mancini, J. A., and Dutton, P. L. (2017). “Making maquette models of bioenergetic structures,” in *Mechanisms of primary energy transduction in Biology*. Editor M. Wikström (London, United Kingdom: The Royal Society of Chemistry), 1–24.
- Munson, M., O'Brien, R., Sturtevant, J. M., and Regan, L. (1994). Redesigning the hydrophobic core of a four-helix-bundle protein. *Protein Sci.* 3 (11), 2015–2022. doi:10.1002/pro.5560031114
- Pace, C. N., and Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* 75 (1), 422–427. doi:10.1016/s0006-3495(98)77529-0
- Page, C. C., Moser, C. C., Chen, X., and Dutton, P. L. (1999). Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* 402 (6757), 47–52. doi:10.1038/46972
- Pasternak, A., Kaplan, J., Lear, J. D., and DeGrado, W. F. (2001). Proton and metal ion-dependent assembly of a model diiron protein. *Protein Sci.* 10 (5), 958–969. doi:10.1110/ps.52101

- Pfeil, W., Nolting, B. O., and Jung, C. (1993). Apocytochrome P450cam is a native protein with some intermediate-like properties. *Biochemistry* 32 (34), 8856–8862. doi:10.1021/bi00085a017
- Pirro, F., Schmidt, N., Lincoff, J., Widel, Z. X., Polizzi, N. F., Liu, L., et al. (2020). Allosteric cooperation in a de novo-designed two-domain protein. *Proc. Natl. Acad. Sci. U. S. A.* 117 (52), 33246–33253. doi:10.1073/pnas.2017062117
- Polizzi, N. F., Wu, Y. B., Lemmin, T., Maxwell, A. M., Zhang, S. Q., Rawson, J., et al. (2017). De novo design of a hyperstable non-natural protein-ligand complex with sub-Å accuracy. *Nat. Chem.* 9 (12), 1157–1164. doi:10.1038/nchem.2846
- Rackham, O. J., Madera, M., Armstrong, C. T., Vincent, T. L., Woolfson, D. N., and Gough, J. (2010). The evolution and structure prediction of coiled coils across all genomes. *J. Mol. Biol.* 403 (3), 480–493. doi:10.1016/j.jmb.2010.08.032
- Raleigh, D. P., and Degrado, W. F. (1992). A denovo designed protein shows a thermally induced transition from a native to a molten globule-like state. *J. Am. Chem. Soc.* 114 (25), 10079–10081. doi:10.1021/ja00051a061
- Rau, H. K., and Haehnel, W. (1998). Design, synthesis, and properties of a novel cytochrome b model. *J. Am. Chem. Soc.* 120 (3), 468–476. doi:10.1021/ja973018r
- Regan, L., and DeGrado, W. F. (1988). Characterization of a helical protein designed from first principles. *Science* 241 (4868), 976–978. doi:10.1126/science.3043666
- Richert, S., Tait, C. E., and Timmel, C. R. (2017). Delocalisation of photoexcited triplet states probed by transient EPR and hyperfine spectroscopy. *J. Magn. Reson.* 280, 103–116. doi:10.1016/j.jmr.2017.01.005
- Robertson, D. E., Farid, R. S., Moser, C. C., Urbauer, J. L., Mulholland, S. E., Pidikiti, R., et al. (1994). Design and synthesis of multi-haem proteins. *Nature* 368 (6470), 425–432. doi:10.1038/368425a0
- Rojas, N. R., Kamtekar, S., Simons, C. T., McLean, J. E., Vogel, K. M., Spiro, T. G., et al. (1997). De novo heme proteins from designed combinatorial libraries. *Protein Sci.* 6 (12), 2512–2524. doi:10.1002/pro.5560061204
- Ruckthong, L., Zastrow, M. L., Stuckey, J. A., and Pecoraro, V. L. (2016). A crystallographic examination of predisposition versus preorganization in de Novo designed metalloproteins. *J. Am. Chem. Soc.* 138 (36), 11979–11988. doi:10.1021/jacs.6b07165
- Sanchez-Baracaldo, P., and Cardona, T. (2020). On the origin of oxygenic photosynthesis and Cyanobacteria. *New Phytol.* 225 (4), 1440–1446. doi:10.1111/nph.16249
- Saribas, A. S., Ding, H., Dutton, P. L., and Daldal, F. (1997). Substitutions at position 146 of cytochrome b affect drastically the properties of heme bL and the Qo site of Rhodospirillum rubrum cytochrome bL complex. *Biochim. Biophys. Acta* 1319 (1), 99–108. doi:10.1016/s0005-2728(96)00120-x
- Schluchter, W. M., Shen, G., Alvey, R. M., Biswas, A., Saunee, N. A., Williams, S. R., et al. (2010). Phycobiliprotein biosynthesis in cyanobacteria: Structure and function of enzymes involved in post-translational modification. *Adv. Exp. Med. Biol.* 675, 211–228. doi:10.1007/978-1-4419-1528-3_12
- Sharp, R. E., Diers, J. R., Bocian, D. F., and Dutton, P. L. (1998). Differential binding of iron(III) and zinc(II) protoporphyrin IX to synthetic four-helix bundles. *J. Am. Chem. Soc.* 120 (28), 7103–7104. doi:10.1021/ja980432y
- Solomon, L. A., Kodali, G., Moser, C. C., and Dutton, P. L. (2014). Engineering the assembly of heme cofactors in man-made proteins. *J. Am. Chem. Soc.* 136 (8), 3192–3199. doi:10.1021/ja411845f
- Soltau, S. R., Niklas, J., Dahlberg, P. D., Poluektov, O. G., Tiede, D. M., Mulfort, K. L., et al. (2015). Aqueous light driven hydrogen production by a Ru-ferredoxin-Co biohybrid. *Chem. Commun.* 51 (53), 10628–10631. doi:10.1039/c5cc03006d
- Sommer, D. J., Vaughn, M. D., Clark, B. C., Tomlin, J., Roy, A., and Ghirlanda, G. (2016). Reengineering cyt b562 for hydrogen production: A facile route to artificial hydrogenases. *Biochim. Biophys. Acta* 1857 (5), 598–603. doi:10.1016/j.bbabi.2015.09.001
- Summa, C. M., Lombardi, A., Lewis, M., and DeGrado, W. F. (1999). Tertiary templates for the design of diiron proteins. *Curr. Opin. Struct. Biol.* 9 (4), 500–508. doi:10.1016/S0959-440X(99)80071-2
- Summa, C. M., Rosenblatt, M. M., Hong, J. K., Lear, J. D., and DeGrado, W. F. (2002). Computational de novo design, and characterization of an A(2)B(2) diiron protein. *J. Mol. Biol.* 321 (5), 923–938. doi:10.1016/s0022-2836(02)00589-2
- Tebo, A., Quaranta, A., Pecoraro, V. L., and Aukauloo, A. (2021). Enhanced photoinduced electron transfer through a tyrosine relay in a De Novo designed protein scaffold bearing a photoredox unit and a Fe^{II}S₄ site. *Chemphotochem* 5 (7), 665–668. doi:10.1002/cptc.202100014
- Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallogr. D. Biol. Crystallogr.* 56 (8), 965–972. doi:10.1107/s0907444900005072
- Torres Martin de Rosales, R., Faiella, M., Farquhar, E., Que, L., Andreozzi, C., Pavone, V., et al. (2010). Spectroscopic and metal-binding properties of DF3: An artificial protein able to accommodate different metal ions. *J. Biol. Inorg. Chem.* 15 (5), 717–728. doi:10.1007/s00775-010-0639-9
- Umena, Y., Kawakami, K., Shen, J. R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* 473 (7345), 55–60. doi:10.1038/nature09913
- Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., et al. (2004). REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D. Biol. Crystallogr.* 60 (1), 2184–2195. doi:10.1107/S0907444904023510
- Vallone, B., Nienhaus, K., Brunori, M., and Nienhaus, G. U. (2004). The structure of murine neuroglobin: Novel pathways for ligand migration and binding. *Proteins* 56 (1), 85–92. doi:10.1002/prot.20113
- Vinyard, D. J., Ananyev, G. M., and Dismukes, G. C. (2013). Photosystem II: The reaction center of oxygenic photosynthesis. *Annu. Rev. Biochem.* 82, 577–606. doi:10.1146/annurev-biochem-070511-100425
- Vinyard, D. J., Badshah, S. L., Riggio, M. R., Kaur, D., Fanguy, A. R., and Gunner, M. R. (2019). Photosystem II oxygen-evolving complex photoassembly displays an inverse H/D solvent isotope effect under chloride-limiting conditions. *Proc. Natl. Acad. Sci. U. S. A.* 116 (38), 18917–18922. doi:10.1073/pnas.1910231116
- Walshaw, J., and Woolfson, D. N. (2001). Socket: A program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* 307 (5), 1427–1450. doi:10.1006/jmbi.2001.4545
- Weber, P. C., Bartsch, R. G., Cusanovich, M. A., Hamlin, R. C., Howard, A., Jordan, S. R., et al. (1980). Structure of cytochrome c: A dimeric, high-spin haem protein. *Nature* 286 (5770), 302–304. doi:10.1038/286302a0
- Willis, M. A., Bishop, B., Regan, L., and Brunger, A. T. (2000). Dramatic structural and thermodynamic consequences of repacking a protein's hydrophobic core. *Structure* 8 (12), 1319–1328. doi:10.1016/s0969-2126(00)00544-x
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* 67 (4), 235–242. doi:10.1107/S0907444910045749
- Woodbury, N. W., Parson, W. W., Gunner, M. R., Prince, R. C., and Dutton, P. L. (1986). Radical-pair energetics and decay mechanisms in reaction centers containing anthraquinones, naphthoquinones or benzoquinones in place of ubiquinone. *Biochim. Biophys. Acta* 851 (1), 6–22. doi:10.1016/0005-2728(86)90243-4
- World Energy Council (2013). “World energy resources: Solar,” in *World energy resources 2013* (World Energy Council). Available at: <https://www.worldenergy.org/publications/entry/world-energy-resources-2013-survey>.
- Wright, C. A., and Clayton, R. K. (1974). The absolute quantum efficiency of bacteriochlorophyll photooxidation in reaction centres of Rhodospirillum rubrum spheroides. *Biochim. Biophys. Acta* 333 (2), 246–260. doi:10.1016/0005-2728(74)90009-7
- Yang, X., Le Brun, N. E., Thomson, A. J., Moore, G. R., and Chasteen, N. D. (2000). The iron oxidation and hydrolysis chemistry of *Escherichia coli* bacterioferritin. *Biochemistry* 39 (16), 4915–4923. doi:10.1021/bi992631f
- Yankovskaya, V., Horsefield, R., Tornroth, S., Luna-Chavez, C., Miyoshi, H., Leger, C., et al. (2003). Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* 299 (5607), 700–704. doi:10.1126/science.1079605
- Zhang, M., Bommer, M., Chatterjee, R., Hussein, R., Yano, J., Dau, H., et al. (2017). Structural insights into the light-driven auto-assembly process of the water-oxidizing Mn4CaO5-cluster in photosystem II. *Elife* 6, e26933. doi:10.7554/elife.26933
- Zhang, Y., and Skolnick, J. (2005). TM-Align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 33 (7), 2302–2309. doi:10.1093/nar/gki524
- Zhuang, J., Amoroso, J. H., Kinloch, R., Dawson, J. H., Baldwin, M. J., and Gibney, B. R. (2004). Design of a five-coordinate heme protein maquette: A spectroscopic model of deoxymyoglobin. *Inorg. Chem.* 43 (26), 8218–8220. doi:10.1021/ic048502r