

Creating Functional Artificial Proteins

Reza Razeghifard^{*‡}, Brett B. Wallace[‡], Ron J. Pace[§] and Tom Wydrzynski[‡]

[‡]Research School of Biological Sciences and [§]Department of Chemistry, The Australian National University, Canberra, ACT 0200, Australia

Abstract: Much is now known about how protein folding occurs, through the sequence analysis of proteins of known folding geometry and the sequence/structural analysis of proteins and their mutants. This has allowed not only the modification of natural proteins but also the construction of *de novo* polypeptides with predictable folding patterns. Structure/function analysis of natural proteins is used to construct derived versions that retain a degree of biological activity. The constructed versions made of either natural or artificial sequences contain critical residues for activity such as receptor binding. In some cases, the functionality is introduced by incorporating binding sites for other elements, such as organic cofactors or transition metals, into the protein scaffold. While these modified proteins can mimic the function of natural proteins, they can also be constructed to have novel activities. Recently engineered photoactive proteins are good examples of such systems in which a light-induced electron transfer can be established in normally light-insensitive proteins. The present review covers some aspects of protein design that have been used to investigate protein receptor binding, cofactor binding and biological electron transfer.

Keywords: Protein design, peptide complexes, receptor binding, cofactor binding, chlorophyll, redox-active residue, electron transfer, reaction center.

INTRODUCTION

Proteins are amino acid sequences that are folded into a unique structure allowing them to perform specific structural or catalytic functions. The diversity of folded protein structures affords an enormous range of biological functions. The folded structure allows functional groups such as the side chains of residues, organic cofactors or redox-active metals to come into close proximity with each other. All residues play a role in the overall structure and stability of the protein, although only a small number of them have a key functional role. These key residues can be involved in enzymatic activity, receptor binding, cofactor binding or redox chemistry. In this review, we summarize some of progress made in creating functional artificial proteins from natural or synthetic amino acid sequences in relation to receptor binding, metal and porphyrin binding, and constructing photoactive proteins.

PROTEIN STRUCTURE

The uniquely configured native structure of a protein is formed from a random coil through a transition where an astronomical number of other configurations can theoretically occur. For proteins with < 100 residues, the transition often appears to be without any kinetic intermediates and involves only two states, native and denatured [1, 2]. The chain length seems to be the main criteria for two-state folding. The stability of proteins with a two-state folding behaviour can change by 2-8 kcal/mole while the unfolding and refolding rates can vary by more than a factor of 10⁵ [2]. The

rate-limiting step in the folding transition is the rate of the formation of a particular conformation in a set known as the transition-state ensemble which is an ensemble of structures with similar energies [3]. In a thermodynamic view, a sufficiently large energy gap is required to favor the native state over the unfolded state and partially folded states. The lowest energy conformation is stabilized by a number of factors including hydrophobic effects, van der Waals packing interactions, hydrogen bonding and disulfide covalent bond formation. The topology of the transition-state ensemble is similar to that of the native state, but it lacks native-like properties such as the full formation of secondary structure, the close packing of side chains and complete burial of hydrophobic residues from the solvent. For example, the study of mimics of partially unfolded intermediates for cyt *b*₅₆₂ at atomic resolutions showed that these intermediates had native-like topology but lacked native-side chain interactions [4].

Proteins that perform the same function often have a different amino acid sequence in different species. In many cases modifying a protein results in no loss of stability and even sometimes produces a stability increase [5]. For example, many studies have been performed to understand the relatively high-temperature stability of thermophilic and hyperthermophilic proteins. The thermostability of these proteins seem to be related to an increased proportion of hydrophobic residues with branched side chains and charged residues as well as deletion or shortening of loops [5]. In a recent study made on cytosine deaminase, an enzyme that converts cytosine to uracil, it was shown that three substitutions of one hydrophobic residue for another (A23L/I140L/V108I) produced a 10 °C increase in apparent melting temperature *T*_m and a 30-fold increase in half-life of the enzyme at 50 °C, with no reduction in its catalytic activity [6]. The crystal

*Address correspondence to this author at the Research School of Biological Sciences, and Department of Chemistry, The Australian National University, Canberra, ACT 0200, Australia; E-mail: reza.razeghifard@anu.edu.au

structure of the mutant at 1.7 Å revealed the tight packing of these residues in the enzyme core filling cavities providing a 70 Å² additional buried surface area. Mutation of surface-exposed residues can also play an important role in protein stability. The enhanced stability of a thermophilic cold shock protein (*Bc-Csp*) compared to its counterpart from a mesophile (*Bs-CspB*) was shown to be mainly due to electrostatic interactions of Arg 3 and hydrophobic contribution of Leu 66 at the C-terminus [7]. When both residues were introduced into *Bs-CspB* the midpoint of the thermal unfolding transition was increased by 20.6 °C. Both *Bs-CspB* and *Bc-Csp* are small (66-67 residues), monomeric proteins with no disulfide bonds.

On the other hand, proteins are also known to exist in an ensemble of conformers with the equilibrium shifting in favor of the native conformation upon binding to their targets (for a review see [7]). In some extreme cases, functionally active proteins are found to exist in natively unfolded conformations. The advantage of molecular disorder seems to be in adopting various folded structures to recognize different biological partners. Some examples of these proteins are the fibronectin-binding domain protein [8], tau protein β [9], prothymosin α [10] and the cyclin-dependent kinase inhibitor p21^{waf1/cip1/sdi1} [11]. In the case of these proteins, the disordered structure and conformational variability are related to their binding diversity that is required for their biological activity. When designing a ligand for a receptor, the important factor is that the actual structural target of the ligand is achieved upon receptor binding. This is often the case for designing short peptides as the ligand because short peptides lack the complexity necessary for forming a stable conformation [12].

The implication is that since natural proteins can tolerate a variation in the sequence, there is a potential for redesign. However, successful design of proteins from scratch is limited by our current knowledge of the relationship between the amino acid sequence and the folded structure of a protein. There has been some recent progress in making small artificial proteins with a desired fold using mainly two approaches, rational and combinatorial designs. In the rational design, a sequence is chosen such that the desired target is achieved based on rules governing folding a sequence into a particular geometry. The rational design often proceeds through an iterative process. The design strategy can be positive, incorporating rules directed towards obtaining the target fold, or negative by including rules disfavoring alternative folds. In the combinatorial approach, variants are selected from libraries of sequences with the specific target structure or activity. Phage display is a common technique for this approach.

PROTEIN DESIGN

One application of protein design is to construct simpler versions of proteins using either artificial or natural sequences. At present, one of the main challenges is to give thermodynamically stable designed proteins native-state characteristics. The native-like properties can be achieved by including sufficient structural complexity in the design. Thermodynamic and kinetic understanding of the direct formation process of a native structure from an unfolded state is

important for *de novo* design to avoid the formation of molten globular states [13]. Compared to the design of β-sheets, the design of an α-helix fold has been more successful. This is because interactions required for the formation and stabilization of an α-helical structure are between local residues while it can be long range for a β-sheet involving residues that are ~20 amino acids apart in sequence. An α-helix is basically formed by the formation of a hydrogen bond between the carbonyl oxygen of residue *n* with the amine hydrogen of residue *n*+4.

An alternating pattern of hydrophobic and hydrophilic residues in a heptad repeat is sufficient for the formation of an α-helical structure [14]. For a heptad repeat, labeled *a-g*, *a* and *d* are typically hydrophobic while *b,c,e,f,g* are usually hydrophilic. In aqueous solution, designed helical peptides self-assemble into oligomeric structures such as a helix bundle in order to shield the hydrophobic residues and give them further stabilization. An example of one such design is the antiparallel four-helix bundle protein formed from the *de novo* design of four monomeric helices (24 residues) using a heptad repeat template [15]. This mimics a commonly occurring motif in natural proteins [16]. However, in general the construction of monomeric helices using a heptad repeat template is not sufficient alone to ensure a native-like structure. It is important that each hydrophobic residue on each helix monomer be selected carefully to maximize the stability of the helix bundle protein. Three rounds of iterative design were used to find the most promising variants carrying Leu, Ile, Val, Ala or Phe in the *a* and *d* positions to produce four-helix bundle proteins with native-like structure [17]. Additional constraints can be placed on the properties of the four-helix bundle proteins by linking the α-helix monomers with loops or disulfide bonds. Four-helix bundle proteins made of two helix-loop-helix [18] or two disulfide-bridged helical peptides [19] were constructed with NMR structures showing both anti or syn topologies. The incorporation of loops between individual α-helices enabled the construction of three-helix proteins and provided a way to control the orientation of helices [20]. Examples of *de novo* designed four-helix bundles where all the helices are linked by loops were also presented [21].

On the other hand, β-hairpins are comprised of two antiparallel β-strands, with a building unit of a polar-nonpolar pair, joined by a loop (for a review see [22]). For example, a three-stranded antiparallel polypeptide was constructed that could cooperatively fold in aqueous solution [23]. Placement of aromatic residues, ionic interactions and the loop sequence all played a role in this successful design. The burial of hydrophobic groups and the interaction between aromatic side-chains seemed to stabilize the three-stranded conformation [24]. Recently, the design and characterization of a homotetrameric miniprotein (BBA) was reported in which each monomer (21 residues) adopted a mixed ββα secondary structure [25]. The compact BBA miniprotein was then redesigned to convert it into a heterotetramer [26]. A successful design of a novel fold, not present in the Protein Structure Data Base, was reported for a 93-residue α/β protein, called Top7 [27]. Top7 was shown to be monomeric with a good thermal stability and its structure was resolved by X-ray crystallography.

1. RECEPTOR BINDING

Protein-protein interactions between receptors and their ligands are generally quite specific. The specificity originates from the spatial arrangement of key contact residues on the ligand and their interaction with partners on the receptor. The interactions are in the form of hydrogen bond and hydrophobic or electrostatic interactions. It seems possible to create an antagonist through protein design by introducing interacting residues on the surface of another protein or a miniprotein. The miniprotein needs to have a stable structure preserving the spatial geometry of the binding interface or acquire such geometry after binding. To successfully design an antagonist, knowledge of the X-ray or NMR solution structure of the receptor-ligand complex is usually required. For example, a mimic of interleukin-4 (IL4) was created by transferring the key residues for binding to the IL4 α receptor (IL4-R α) onto helices of GCN4 [28] or DHP [29] (Fig. 1). While GCN4 is a natural two α -helical protein, DHP is a *de novo* designed four α -helix protein. In the case of GCN4, the antagonist was made by introducing 8 residues of IL4 and a disulfide bridge cross-linking two helices to improve binding affinity and stability. For DHP, an improved antagonist was achieved using 14 residues of IL4 even though its affinity for IL4-R α is still 2.7×10^4 times lower than that of IL4. A loss in stability by 5 kcal/mol was seen for this new designed DHP protein, but its structure was still well-ordered leading to an instructive crystal structure.

In another study, to create a fully α -helical and biologically active version of a C-peptide (C34), the complete binding epitope (19 residues) was grafted onto the surface of the GCN4 protein (C34-GCN4) [30]. C-peptides, which are derived from the C-terminal region of HIV-1 gp41, have an unstructured conformation when unbound but are α -helical when bound to the gp-41 N-peptide region inhibiting HIV-1 entry into cells. To improve the poor helical content and solubility of the C34-GCN4, it was linked by a disulfide-bridge to the GCN4 protein forming a heterodimer denoted C34 coil. C34 coil exhibited potent antiviral activities with IC₅₀ values of some nM which is an order of magnitude better than those of C34. C34 coil was also shown to be about 1,000 fold more resistance to proteolytic degradation. A 13-mer peptide (WRYESSLEPYPD) was designed to bind to snake-venom α -bungarotoxin with high affinity (IC₅₀ of 2 nM) inhibiting the toxin interactions with the nicotinic acetylcholine receptor [31]. This peptide was derived from a previous version, which was isolated using a phage display library, based on structural information of the peptide-toxin complex. The peptide appears to fold into a β -hairpin structure created by two antiparallel β -strands.

A mimic of interleukin-5 (IL5) was identified by random mutagenesis of the turn in a coiled coil miniprotein (CCSL) using the phage display technique [32]. IL5 is a homodimer comprised of two α -helical bundles [33]. The structure also contains two antiparallel β -sheet motifs on opposite sides of the molecule. CCSL is a 56-residue polypeptide with a disulfide bond connecting the N- and C- terminus [34]. It appears that the dominant selectant was the CCSL variant with a turn sequence (PVEGRV) resembling the β -strand sequence in the CD turn region of IL5. For its binding ability, this variant of CCSL contains a Glu/Arg pair similar to Glu89/Arg91 of

IL5 which are believed to be crucial for binding to IL5 α receptor. The phage display technique was also used to search for a peptide antagonist of interleukin-1 (IL1) [35]. Some of these peptides were able to block binding of IL1 α to the type I IL1 receptor with IC₅₀ values of about 2 nM. The binding was shown to be selective since these peptides did not bind the human type II IL1 receptor or the murine type I IL1 receptor. A mini-antagonist peptide (14-residues) based on the loop AB of IL6 was synthesized by computer-aided design [36]. The peptide was shown to compete with IL6 for binding to human IL6 α receptor.

A 27-amino acid CD24 mimic (2.9 kDa) was rationally designed to inhibit HIV-1 infection by binding to gp120 with a CD4 native-like binding affinity (IC₅₀ of some nM) [37, 38]. This mini-CD4 possesses the ability to unmask conserved neutralization epitopes of gp120 that are cryptic on the unbound glycoprotein. The rational design of this potent mini-CD4 was based on modifying the original CD4M9 miniprotein [39] by computer modelling using the crystallographic structure of the gp120-CD4 [40] as a template. The N-terminal residue of CD4M9 was replaced by thiopropionic acid to prevent unfavorable steric contacts at its N terminus upon binding to gp120. Also Phe23 was replaced by a biphenylalanine to increase its hydrophobic interaction with the apolar Phe43 cavity. The binding affinity was further improved by introducing three more mutations. CD4M9 is derived from the scorpion scyllatotoxin with abolished toxic activity. It contains a solvent exposed β -hairpin similar to that of the CDR2-like region of human CD4 [41].

2. COFACTOR BINDING

It is estimated that more than 30% of all natural proteins require cofactors such as metal ions and organic moieties to carry out their biological functions [42]. Native scaffolds contain crafted binding sites that are proven to have high tolerance for cofactor substitution and mutation. The capacity of binding sites to tolerate cofactor substitutions has allowed the introduction of novel functions into a native protein without loss of the protein structural integrity. Cofactors can either be covalently bound to the protein backbone or coordinated to residues. There are various groups in proteins that are involved in cofactor ligation: amides from Asn, Gln or the protein chain, aminos from Lys or the N-terminus, carboxyls from Asp, Glu or C-terminus, hydroxyls from Ser or Thr, imidazole from His, sulphide from Met, thiol from Cys and carbonyls from the protein backbone [43]. Examples of cofactors are porphyrins like heme, (bacterio)chlorophyll ((B)Chl), (bacterio)pheophytin ((B)Pheo), and redox-active metals.

2.1 Porphyrins

Various oxidoreductase proteins use porphyrin type cofactors. These cofactors include metalloporphyrins such as heme and Chl. The common heme cofactor is iron protoporphyrin IX. Chls are chlorin tetrapyrroles with a fifth ring fused to one of the pyrrole rings and Mg or Zn as their central metal.

2.1.1. Heme

Heme is used for a variety of functions by proteins. For example, while the hemes in the cyt *b₆f* complex are in-

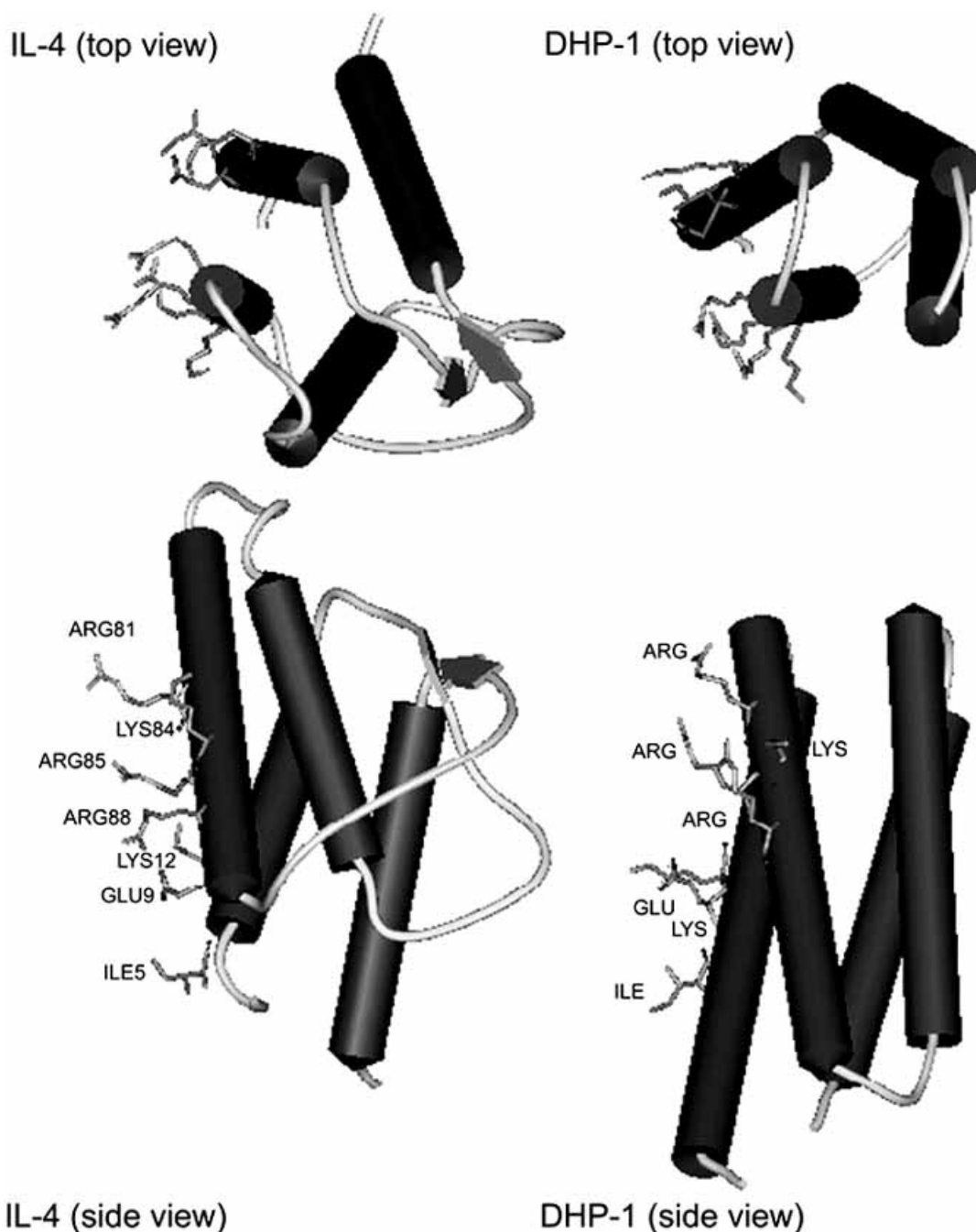


Fig. (1). The structure of IL4 and DHP-14-AB with binding residues for IL4-R α highlighted. IL4 is a 129 amino acid protein containing three disulfide bonds. The IL4 structure is a highly compact left-handed antiparallel four-helix bundle with a short two-stranded antiparallel β -sheet [161, 162]. The helices are labeled A through D for residues 4-19, 40-60, 70-94, and 109-126, respectively. IL4 is a highly dynamic protein with an unfolding free energy of ~ 4 kcal/mol and a denatured-dependent m value ~ 1.3 kcal/mol² [163]. The multiple biological functions of IL4 are mediated through the formation of an oligomeric complex with a heterodimeric receptor complex consists of IL4-R α and either the common γ receptor or the IL-13 receptor $\alpha 1$ chain. The high affinity binding of IL4 to IL4-R α is mediated through the complementary charged residues at the binding interface. The binding interface, determined from the crystal structure of the complex, is located on helices A and C with Glu9 and Arg88 as the most important binding residues [164]. The interaction of Glu9 with the receptor appears to have polar steering effect in the receptor-complex formation [165]. DHP-14-AB represents a *de novo* designed protein with 108 residues carrying binding residues for IL4 α receptor on the adjacent helices A and B. It is a right-handed helical bundle that was created through two generations of design from DHP-1, based on the incorporation of positive elements and avoidance of negative elements. The DHP-14-AB protein showed a cooperative unfolding transition with an unfolding free energy of 8.1 kcal/mol and a well-ordered crystal structure. The low apparent affinity DHP-14-AB for IL4R α , IC₅₀ of 27 μ M, may be due to transplantation of only a subset of IL4 binding residues as well as the geometry of the helices interacting with IL4R α . Figures were generated from PDB entries 1RCB and 1Y4C using DS Viewer (Accelrys, CA) program.

volved in electron transfer (ET), the primary function of heme in myoglobin (Mb) is O₂ storage and transport. The biological function of heme is tuned by the protein environment through factors such as axial ligation, hydrogen bonding, out-of-plane distortion and covalent attachment. For example, by random mutation of the protein environment surrounding the heme in horseheart Mb, the peroxidase activity was significantly increased [44]. Variants of horseheart Mb were screened by separating *E. coli* colonies which turned green as a result of the peroxidase activity. One variant exhibited a 25-fold higher peroxidase activity compared to the wild-type Mb. In another study, variants of cyt P450 were co-expressed with horseradish peroxidase in *E. coli* to screen for those with improved hydroxylation reactivity [45]. The peroxidase enzyme was used to convert the product of cyt P450 into a fluorescent compound for the activity assay.

The redox potential of heme in natural proteins is highly variable and spans over 800 mV in range (from +400 mV for cyt *b*₅₅₉ to -400 mV for cyt *c*3) [46]. For example, the midpoint potential of cyt *c*₅₅₃ which performs anaerobic sulphate respiration is +20 mV [47]. This potential is lower, by 200-450 mV, than for the general cyt *c* family and has been related to a high solvent accessibility of the heme propionate groups [48]. In *c*-type cyts, the heme has an additional covalent attachment, compared to other hemes that are bis-ligated to the protein. For cyt *b*₅₆₂ when the ligation of the heme was mutated from the native His/Met to the classic cyt *b*-type bis-His ligation, the midpoint potential at pH 7 decreased by 270 mV and the pH dependence of the *E*_m was considerably al-

tered [49]. A new type of heme, named heme *x*, was recently discovered in the cyt *b*_{6f} complex [50]. This atypical heme is covalently bound to the protein but lacks any protein axial ligands. The axial water ligand of heme *x* is hydrogen bonded to the propionate group of an adjacent heme. The function of heme *x* is currently unclear.

De novo designed heme binding proteins also exhibit catalytic activity such as CO binding, peroxidase activity and ET reactions. Eight of these proteins derived from combinatorial libraries were studied to measure their affinity for CO, the kinetics of CO binding and release, and to analyze the local hydrogen-network environment around the bound CO [51]. The CO binding affinities were similar to that of engineered Mb, with a narrow range of dissociation rate constants from 0.028 to 0.11 s⁻¹. For natural heme proteins, the CO dissociation rate constant can vary from 7.2 × 10⁻⁵ s⁻¹ for horseradish peroxidase to 6.5 s⁻¹ for cyt P450. When synthetic heme proteins were screened for the peroxidase activity, one was found to have a turnover number of 17000 min⁻¹ that is fourfold lower than horseradish peroxidase but higher than μperoxidase [52]. Modulation of the redox potential was also achieved in simple heme maquettes [46]. The redox potential was varied over 435 mV through altering the electron-donating/withdrawing nature of the peripheral macrocycle substituents of the heme and the protonation/deprotonation state of neighboring charged residues. Some of the activity differences between natural and designed proteins are due to partial shielding of the heme in

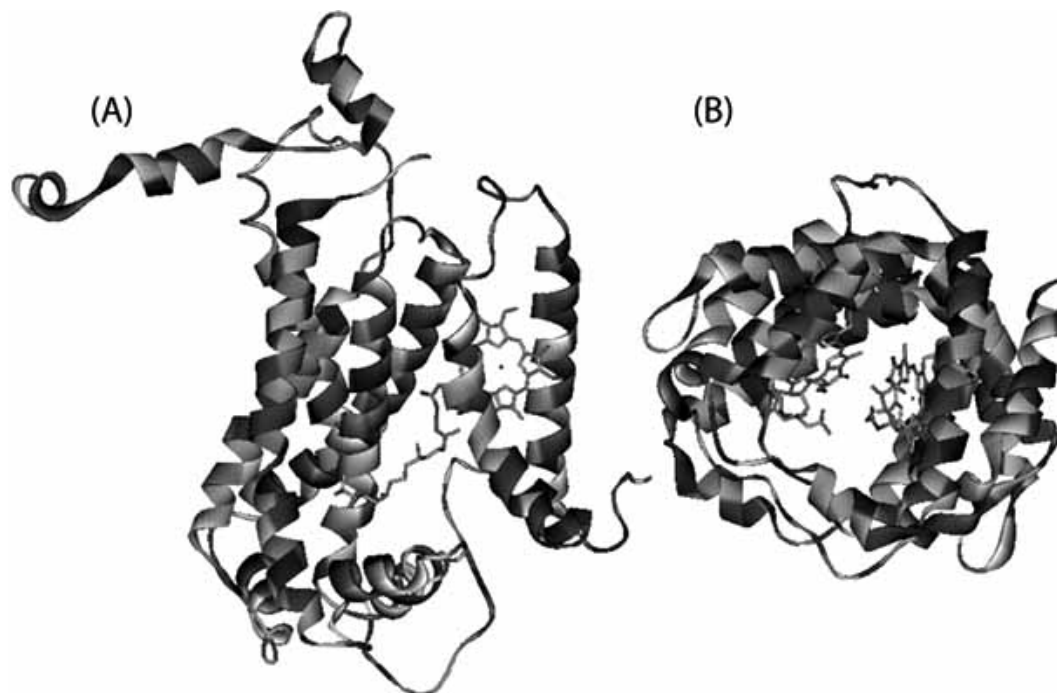


Fig. (2). Chl binding sites in the cyt *b*_{6f} complex (A) and peridinin-Chl protein (B). The Chl *a* in cyt *b*_{6f} is a single Chl molecule ligated to the protein possibly by the carbonyl group of a Gly with an intervening water molecule. The phytol chain ends within the Q_o pocket where hydroplastoquinone oxidation and proton release take place. Less bulky side chains at the opposite wall of the cleft compensate for the presence of the phytol chain. The Chl *a* function in cyt *b*_{6f} is not yet understood. The peridinin-chlorophyll *a*-protein (PCP) acts as an accessory photosynthetic light-harvesting pigment-protein complex in the dinoflagellate *Amphidinium carterae*. The complex transfers its excitation energy to PSII. The distance between the centers of the two Chl *a* molecules in the 30.2-kDa monomer of PCP is 17.4 Å. Peridinin pigments are not shown. Figures were generated from PDB entries 1PPR and 1Q90 using DS Viewer (Accelrys, CA) program.

The cofactor may also perform a structural role that is important for the stability of the holo-protein. For example, in the absence of heme, apo cyt b_{562} is a molten globule due to partial unfolding of helix IV and a poor packing interface between helix III and IV [53]. The stability is given by the heme cofactor, which has an area of 1.0 nm^2 , by providing crucial hydrophobic contacts. To convert apo cyt b_{562} into a more stable structure, the phage display technique coupled with proteolysis was used to redesign this protein [53]. The rationale was to increase the stability by filling the heme-binding pocket cavity. Three residues of Arg98, His102 and Arg 102 (RHR) were randomly mutated to accomplish this. The axial ligand of the heme, Met 7, was also changed to a bulky Trp residue. The structure of two of the most stable variants ING and IRL, which surprisingly used a charged residue in the place of His, were then resolved by NMR or X-ray crystallography. This showed that all four helices were formed in these variants. The heme is also thought to assist the rapid submillisecond folding of cyt b_{562} by reducing the conformational freedom of the polypeptide chain. The assumption is that the heme is still bound to the denatured state of cyt b_{562} [54] or the axial ligation of the heme Fe(II) to Met is the first step in the heme-facilitated refolding of cyt c - b_{562} and cyt c_{556} [55]. Cyt c - b_{562} is a variant of cyt b_{562} in which heme is covalently linked to the protein like c -type cytochromes.

Like natural proteins, heme can also play a structural role for *de novo* designed proteins. A four helix bundle protein made of four monomeric peptides was designed to selectively bind two Fe(III)-diphenylporphyrin (DPP-Fe(III)) cofactors through *bis*-His ligation [56]. In the absence of DPP-Fe(III), the protein was mainly unstructured but the assembled DPP-Fe(III)-protein complex gave rise to a well-dispersed 1-D proton NMR spectrum which is indicative of a native-like structure. The structural specificity was cofactor dependent since the complex of heme with this protein lacked a well defined structure, as shown by its CD spectra. A successful design of another native-like complex was also reported for a di-heme four-helix bundle protein which was constructed as a model for cyt bc_1 [57, 58].

2.1.2. Chlorophyll

(B)Chl pigments are the photocatalysts in membrane-bound photosynthetic protein complexes. These large pigment-protein complexes have evolved into multi-component systems, which efficiently capture light energy and convert it into an electron flow. The light energy captured by (B)Chls in the light-harvesting systems is channeled to special photo-oxidizable (B)Chls in the reaction centre (RC) where the photochemistry occurs. These special photo-oxidizable (B)Chls are called primary electron donors (P) and are the key determinant for the unique function of each photosystem. Two photosystems (PSI and PSII) work in tandem in plants and cyanobacteria while only one photosystem (the bacterial reaction centre) is found in anoxygenic photosynthetic bacteria. Even though fundamental principles of the photochemical reactions are conserved, the RC of each photosystem is unique according to the chemical nature of its electron acceptor and donor [59]. All known photosynthetic RCs have two branches of structurally parallel redox cofac-

tors and pathways, but generally only one branch is used for functional ET.

The interaction of Chl with the binding protein is complex because of the large hydrophobic phytyl tail of Chl. The binding sites for Chls in natural proteins are tailored so that both the entire porphyrin head and the phytyl tail are buried within the protein making contact with distal residues (Fig. 2). Crystallographic determinations of antenna and RC proteins have unveiled a variety and multiplicity of structural configurations [60, 61, 62, 63, 64, 65, 66, 67]. There are several types of Chls (*a-d*) that are different only in their peripheral groups. For example, Chl *b* and *d* are structurally identical to Chl *a* except for the substitution of the methyl to a formyl group for Chl *b* and the vinyl to a formyl group in Chl *d*. The optical spectra are however significantly different. Despite the small structural differences, Chl binding sites in proteins can be very selective. In addition, PSII and the bacterial RC contain two (B)Pheo molecules which are (B)Chl molecules the central metal is replaced by two protons. They are the primary electron acceptors which participate in transferring electrons from the excited P to a tightly bound quinone on the acceptor side of the photosystem.

If the affinity of Chl molecules for the apo-protein is mainly due to hydrophobic interactions between the porphyrin head and residue side chains, then what controls the binding selectivity? The central metal of Chl molecules are penta-coordinated with four ligands provided by the tetrapyrrole nitrogens. The fifth, axial ligand is usually the side chain of an amino acid such as His, Glu, Asn, Gln, Ser, Met, Gly although water molecules and phospholipids are also known to serve as axial ligands. For the recruitment of Chl or Pheo, the decisive factor seems to be the presence or absence of a ligating residue. For the bacterial RC, introduction of M-H202L and L-H173L mutations causes the replacement of two BChls with two BPheo molecules, while introduction of the M-L214H mutation favors the incorporation of a BChl into the BPheo binding site [68]. By introducing a His in place of Leu (D1-L210H), it was shown that the inactive Pheo in the RC of PSII was replaced with a Chl impairing the charge separation capacity [69]. The forward ET was substantially reduced giving rise to photoaccumulation of reduced Pheo. The binding site specificity for a particular Chl type can be related to the nature of the ligating residue. For example, mutagenesis studies of the CP29 antenna protein in PSII showed that glutamine is the favored ligand for Chl *b* while glutamate favors Chl *a* [70]. The CP29 variants were over-expressed in bacteria and reconstituted with purified Chl molecules *in vitro*. This complicated binding mode makes the design of small proteins for Chl binding very challenging.

There is also another class of Chl-containing proteins that are water soluble, unlike photosystems. Examples of these proteins are water-soluble Chl binding proteins (WSCP) [71] and the peridinin-Chl protein [72] (Fig. 2). The peridinin-Chl protein is a trimer in which each monomer consists of two α -helical domains. The pigments are organized into two clusters, each containing one Chl and four peridinin molecules and are positioned in the hydrophobic space between the domains. The distance between the Chls of the two clusters is 17.4 \AA . However, the structures of WSCPs are yet to be

determined. Biochemical studies of WSCPs showed that these proteins form tetramers upon binding Chl. The number of bound Chl molecules can vary from 1-4 depending on species. The tetrameric form exists only in the presence of Chl indicating that the phytol tail is shielded from the solvent upon oligomerization [73]. Other Chl derivatives such as Zn-pheophorbide *a* or chlorophyllide *a* and *b* can bind to the WSCP protein but do not make oligomers. Binding appears to be driven by axial ligation since Pheo cannot bind to this protein.

The binding of Chl molecules to designed peptides, however, requires the presence of detergent to increase the solubility of hydrophobic Chl in aqueous solutions. The binding is usually performed in the presence of *n*-octyl- β -D-glycopyranoside or β -dodecyl-maltoside micelles. Like the natural WSCPs, Chl-peptide complexes oligomerize forming large complexes upon binding Chl molecules resembling light harvesting (LH) complexes. Artificial LH complexes were created by binding BChl *a* molecules to synthetic four α -helical peptides (57 residues) containing either two or four His residues [74]. Complex formation was verified by a large red-shift (777 to 863 nm) in the Q_y absorption band of Chl. A clear red-shift was not observed for binding BChl *a* to the synthetic four α -helical peptides with 4 His residues perhaps due to its poor hydrophobic packing. Binding studies were also performed with Zn-BChl *a*, since BChl *a* was easily decomposed under the experimental conditions used. In another study, a pigment-peptide complex containing a Chl dimer was constructed by adding Chl dissolved in THF to the peptide containing a His residue [75]. The binding was completed after 22 h and was shown to be His dependent, as the control peptide appears to form aggregates with Chls. In another approach, segments of amino acid sequence of the bacterial core LH complex (LH1) were chemically synthesized to determine structural features required for making natural-like subunit complexes [76]. Truncated peptides that were shorter than the natural LH1 β -polypeptides by 17 to 23 residues were fully competent in binding BChl *a* but did not interact with α -polypeptides to form LH1 type complexes. The formation of a LH1 like complex was only achieved with longer synthetic polypeptides [77]. Another membrane-spanning helical peptide (16 residues) containing a conserved motif of LH complexes was also shown to bind Chl molecules [78]. The specific binding of Chl to the peptide was detergent dependent. A peptide lacking ligation residues (His or Glu) was used as a control to test the non-specific binding of Chl. This approach is generally required as Chls and some peptides can non-specifically aggregate in detergent solubilized systems. Ni-BChl was shown to ligate to His residues in the lipophilic (LP) domain of a novel, designed amphiphilic protein maquette (AP3) [79, 80]. Ni-BChl-AP3 is a four-helix bundle formed from 42-mer peptides co-assembled in detergent micelles. The LP domain (24 residues), which is drawn from a *de novo* designed membrane channel, was joined through a flexible linking sequence to a hydrophilic sequence taken from their previously designed water-soluble maquettes.

2.2. Redox Active Metals

There are two types of metal binding sites in proteins. The first type is a pre-organized binding site while in the

second case the binding site is structured after the metal is bound.

The coordination environment of a metal can be determined by analyzing the structure of natural proteins [81]. Even though metals in proteins have structural, catalytic, and regulatory functions, our focus in this review is on redox-active metals. The activity of a metal in the binding site is affected by the metal core, ligand groups, and second shell interactions. Effects of long-range electrostatic interactions due to the distribution of charged residues throughout the protein and solvent shielding are also important factors. For example, second shell interactions are generally responsible for making the midpoint potential (E_m) value for the Cu(II) center vary in the range from +200 to +1000 mV in different proteins [82]. Even though a similar coordination pattern is present in both rusticyanin and plastocyanin, the redox potential of the Cu(II) in the latter protein is much lower. For blue copper, the second shell appears to include some solvent-exposed residues while for Mn ions the second shell residues are mostly buried. Interestingly, unlike Cu ions that are never ligated by acidic residues, the primary ligands for Mn ions are mostly Glu or Asp residues.

Binding of a metal to a peptide does not appear to require a complex design, however accommodating a metal binding site inside a protein to mimic a biological function is a more challenging task. For example, a single tripeptide sequence Gly-Gly-His can effectively form a tetradentate complex with Cu(II) [83]. Interestingly such a peptide is selective for Cu(II) with minimal binding affinity for other metals such as Mn(II), Fe(II), Co(II), Ni(II), Zn (II), Cd(II), Mg(II) and Ca(II). In terms of *de novo* design, a metal binding site can be introduced for the purpose of improving structural stability as well as introducing function. For synthetic peptides with partially folded conformations, the conformation is often stabilized upon binding of the metal. For example, Ghadiri et al showed that by introducing His or Cys residues in positions *i* and *i* + 4, two small 15 mer-peptides interacted with transition metals to form a bidentate complex [84]. The interaction generated sufficient conformational constraints to turn the peptides into α -helical structures. The metal binding was selective since the peptide carrying the Cys residue binds Cd (II) but not Zn(II). However, introducing function requires an understanding of the second-shell interactions which influence metal binding affinity as well as redox activity. The residues involved in second shell interactions can form hydrogen bonds with the ligands as well as affecting the pKa values of ligands. The design of a metal binding site can be quite simple when all the ligands are contained in a small sequence of amino acids. For example, a tetrahedrally distorted square plane with a redox potential of +75 mV was created by including all four blue Cu(II) ligands into a 22-residue peptide [85]. The peptide (YCSPHQGAGMVGK) is the natural binding loop of plastocyanin containing Cys, Met, and His ligands with a flexible Gly linker to add the fourth ligand. In plastocyanin, copper is ligated to the protein by two His nitrogen atoms, a Cys sulfur donor, and a Met sulfur atom at a longer distance in a distorted tetrahedral geometry [86]. The Cu ion of plastocyanin acts as the electron donor to the oxidized P^+ in PSI. When a His residue is used as the ligand, the metal is ligated through N(ϵ) but the existence of free N(δ) offers an opportunity to introduce sec-

ond shell interactions. For *de novo* designed proteins, this property was used to enforce the perpendicular orientation of the ligating His residues to the porphyrin by creating hydrogen bonds between His N(δ) and Thr O(γ) [58, 87].

2.2.1. Iron-Sulfur Cluster

Iron-sulfur (Fe-S) clusters are iron atoms which are wholly or predominantly coordinated by sulfur atoms usually provided by Cys residues. The E_m of these clusters in proteins varies from +450 mV to -700 mV, which is believed to be controlled by solvent accessibility, hydrogen bonding and the distribution of neighboring charged residues [88]. There are three 4Fe-4S clusters (F_X , F_A and F_B) in the RC of PSI that function as secondary electron acceptors. They are sequentially reduced by an electron transferred from the excited P, via a monomeric Chl molecule and phyloquinone. The electron is finally transferred to the ferredoxin protein (in some cases flavodoxin), which is used to reduce NADP⁺ to NADPH₂, the reducing compound needed for CO₂ fixation. Different methods have been used to incorporate Fe-S clusters into proteins. For example, Gibney et al incorporated two 4Fe-4S clusters into a synthetic four α -helix bundle [89]. Each dimeric peptide contains two helical regions of ~ 27 residues with a *Clostridial* ferredoxin binding domain for the cluster (CEGGCIACGAC) inserted into the glycine loop connecting the helices. The self-assembly of the 4Fe-4S cluster with the peptide was achieved by adding a solution of ferric chloride and sodium sulfide to the chemically reduced peptide under anaerobic conditions. The formation of the 4Fe-4S cluster was confirmed by EPR spectroscopy and its reduction potential was determined to be -350 mV. Further studies on this design were made to determine the most important amino acids in the binding domain for the cluster formation [90, 91]. A similar approach was applied to form a single 4Fe-4S cluster resembling the F_X cluster of PSI by inserting the 10-residue CDGPGRGGTC sequence into interhelical loops 1 and 3 of the $\alpha 4$ protein which was originally designed to form a synthetic four-helix bundle protein [92, 93]. The cluster formation in the holoprotein was confirmed by EPR and shown to have an E_m of -422 mV, which is more positive than that of the natural counterpart. Binding sites for a 4Fe-4S cluster were also included in the design of a four α -helical protein formed by two helix-loop-helix peptides (63 residues), as a mimic of the A-cluster of carbon monoxide dehydrogenase [94]. One cluster was incorporated into each 63 mer-peptide when the peptide was reacted with a preformed cluster $[\text{Fe}_4\text{S}_4(\text{SCH}_2\text{CH}_2\text{OH})_4]^{2-}$ and Ni (II). The loop contained the ferredoxin consensus sequence Cys-Ile-Ala-Cys-Gly-Ala-Cys to bind the 4Fe-4S cluster with a fourth Cys serving as the bridging ligand between the cluster and Ni(II). A miniferredoxin protein was also created from a 31-residue peptide carrying a 4Fe-4S cluster with a redox potential of -370 mV [95]. Recently, a synthetic 4-helix bundle protein containing a tetranuclear Cu₄S₄ cofactor with a strong room temperature luminescence was reported [96]. The folding of the apo-peptide, Ac-K(IEALEGK)₂ (CEACEGK) (IEALEGK) GGY-amide, was induced upon binding of Cu(I) to the Cys-X-X-Cys binding domain.

2.2.2. Manganese Cluster

Multi-Mn clusters are versatile cofactors found in enzymes. Examples of these enzymes are arginase, ribonucleo-

tide reductase, catalase and photosynthetic water oxidase [97]. These proteins are often multimeric, but based on a four-helix bundle monomer motif (e.g. catalase). A common feature found in most binuclear Mn centers is the presence of one or two μ -carboxo bridges derived from Asp or Glu side chains of the protein (Fig. 3). Bridging carboxylates probably serve a functional role beyond merely that of a passive structural bridge to bring the metal ions together. Their size and negative charge enables them to spatially separate and electrically screen the metal ions so that the degree of inter-metallic electronic coupling is small. Catalases catalyze disproportionation of hydrogen peroxide into O₂ and water. Unlike catalases with a heme cofactor, the type II catalases are four-helix bundle proteins containing a dinuclear Mn active site that forms a two-electron catalytic cycle [98]. The catalytic site in the water oxidase contains up to 4 Mn and 1 Ca in a compact, exchange coupled cluster (Fig. 3). It executes the most chemically demanding reaction in nature, anodic oxidation of water to O₂ and protons. This is the kinetically limiting step in the electrolytic decomposition of water into H₂ and O₂. The water splitting site in PSII performs this reaction at close to thermodynamically limiting efficiency (< 0.2V over-voltage), at a high turnover rate ($\sim 10^3 \text{ s}^{-1}$), under mild external pH and in the presence of significant concentrations of environmentally common anions, such as Cl⁻ [99]. Although the full structural detail of the site is yet to be resolved, mutagenesis studies [100] suggest that most of the protein ligands which define the cluster geometry are located in a very small region (~ 15 residues) near the C-terminus of the D1 polypeptide of the PSII reaction centre. This is most unusual for a protein redox center and presumably reflects aspects of the unique chemistry which the metal site executes. In particular, a certain flexibility or lability of the enfolding peptide may be necessary for full assembly and function. Fig. 3D compares the three published medium resolution structures of the PSII D1 protein folding near the luminal region containing the Mn/Ca cluster [60, 63, 64]. Significant variation in the C-terminal structure is evident. At present the implication of this is unclear, as it is not known which of any of the four quasi-stable oxidation states or assembly forms of the site the crystal structures represent. However the possibility is raised that functioning catalytic site analogs may be assembled from small, model peptides [101].

To create a binding site for a dimetal, the due-ferri-1 (DF1) protein was designed that is a dimer formed from two helix-loop-helix proteins [87]. DF1 was shown to bind a di-Mn cluster in the hydrophobic core (Fig. 3) [102]. The two Mn(II) ions are ligated by Glu and His residues with a DMSO molecule, from the crystallization buffer, appearing to form a mono oxo bridge with the dimer. DMSO was included in the buffering system to increase water solubility. When the binding site was further expanded by introducing a Gly in the place of Leu, two Mn ions appeared to be coordinated to two terminal water molecules or bridged by a single water molecule [103]. Another series of DF1 proteins called DFtet proteins were designed later to introduce catalytic activity [104]. Unlike DF1, the four helices in DFtet are distinct chains that are self-assembled in aqueous solution by non-covalent interactions. The binding pocket for a substrate, 4-amiophenol, was then sculpted and optimized by removing

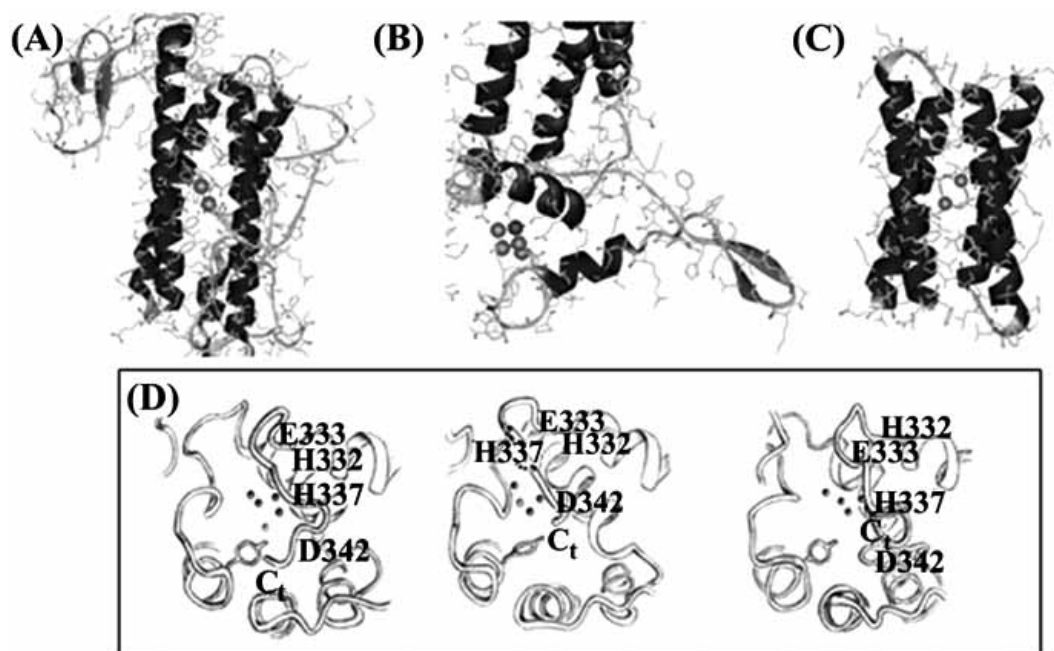


Fig. (3). The structure of Mn clusters in some Mn redox proteins: (A) Mn catalase monomer, (B) PSII and (C) a *de novo* designed DF1 protein. In the type II catalases from mesophilic bacteria or thermophiles, the dinuclear site is located in a 4 helix bundle which multimerizes with some β sheet interactions between monomers. The two Mn ions are μ -oxo, hydroxo and carboxylate, bridged and operate between redox states II-II and III-III in cyclic turnover. In the currently most resolved structure of PSII [64], the Mn ions are proposed to form a cubane-like Mn_3CaO_4 cluster connected to the fourth Mn in the extended region. The metals are largely shielded from solvent by extrinsic proteins whose loss destabilizes the cluster. In the fully reduced form, the cluster contains either Mn(II) or Mn(III) and is progressively oxidized to higher levels (Mn(IV), possibly Mn(V)) during 4 electron turnovers. In the *de novo* designed DF1 protein, the coordination environments of Mn ions were different giving rise to four crystallographically independent Mn-dimers. In three of the four dimers a solvent molecule bridges the two Mn ions (probably water), while in the fourth two terminal solvent molecules are coordinated to the two Mn ions. Based on the coordination geometries, Mn ions are in the Mn(II) oxidation state. The Mn-Mn distance in the solvent-bridged dimers was 3.59 Å while an increase to 4.20 Å was seen for the dimer with two terminal solvent molecules. (D) The D1 peptide folding near the luminal region from the three available medium resolution crystal structures of cyanobacterial PS II. Cartoons were drawn based on the data of Ferreira et al [64], Biesiadka et al [60] and Kamiya and Shen [63]. Residues suggested as ligands from mutagenesis studies [100] are indicated, including the C-terminal (C_t) carboxylate of Ala344. No existing structure locates all of these within binding range of the metal cluster. A progressive C-terminus 'unfolding' may be evident when comparing these structures. Figures were generated from PDB entries 1JKU, 1S5L and 1LT1 using DS Viewer (Accelrys, CA) program.

Leu15 and Ala-19 (corresponding to positions 9 and 13 of DF1, respectively) to avoid unfavorable contacts with the substrate. The variant was capable of catalyzing the two-oxidation of 4-aminophenol to its quinone mono-imine derivative in atmospheric O_2 .

3. REDOX-ACTIVE RESIDUES

Radicals of some amino acids formed from Tyr, Trp, Cys and Gly are known to take part in the enzymatic activities of some metalloenzymes. These residues are involved in ET and can be linked to hydrogen transfer. There are two redox-active Tyr residues, called Y_Z and Y_D , in the PSII protein [105, 106]. The Tyr radicals are paramagnetic and give rise to a structured EPR signal with the line shape dependent on the location of the spin density and the dihedral angle (θ) at the C_β - C_1 bond. The exact function of Y_D is not known but the Y_Z radical is directly involved in water oxidation chemistry [107]. Two water molecules are oxidized producing four electrons, four protons and a molecular oxygen. The oxidation reaction requires four oxidizing equivalents that are sequentially stored in the Mn cluster. The Mn cluster is oxi-

dized by P^+ via Y_Z as the electron carrier. The oxidation rate of the Mn cluster transition can then be obtained by measuring the Y_Z^\bullet re-reduction rate using time-solved EPR [108, 109]. These rates are good indicators of the intactness of the water oxidizing enzyme and are slowed in detergent-treated PSII preparations and when PSII centers are genetically modified [110, 111]. The rates are also affected when the reaction conditions are displaced from optimal, such as by changing pH. This provided mechanistic insights into the enzyme function [112].

Cyt *c* oxidase which reduces O_2 to water uses a cross-linked His-Tyr pair to provide an extra reducing equivalent in the active site, resulting in formation of the oxidized radical species intermediate [113, 114]. The reduction reaction requires three more electrons two coming from the oxidation of heme (Fe^{2+} to Fe^{4+}) and one from the oxidation of Cu^+ to Cu^{2+} . In proposed to form galactose oxidase, a Tyr residue cross-linked to Cys through a thioether bond is used to catalyze the two-electron oxidation of a primary alcohol to the corresponding aldehyde (for a review see [115]). Ribonucleotide reductases (RNR) are involved in the synthesis of

deoxyribonucleotides from their ribonucleotide precursors. Three classes of RNR differing in their metal cofactors have been identified. Class I, Class II and Class III RNR enzymes use a diiron cluster, a cobalt containing cobalamin cofactor or a 4Fe-4S cluster coupled to S-adenosylmethionine, respectively. However, a new R2-type protein from *Chlamydia trachomat* has recently been reported that lacks the tyrosyl radical site and the protein yields an iron-coupled radical instead [116]. A common feature of all RNRs is a conserved Cys residue that is believed to be converted into a thiyl radical in all three classes. The thiyl radical that initiates the substrate turnover is generated by radical transfer from a tyrosyl or glycy radical in Class I and III RNR enzymes, respectively. Transfer of the electron hole occurs over a long distance since Tyr122 on the R2 subunit and Cys439 of the R1 subunit are 35 Å apart. One possible mechanism to describe such long distance ET is a hopping mechanism, where charge is relayed along on a cascade of suitable sites which are electronically coupled to each other [117, 118]. Such a multistep electron-hole transport pathway involving three Trp residues was proposed for the photoactivation process in photolyase enzyme [119].

A 20-mer peptide derived from the C-terminal of the R2 subunit containing a Trp as a phototrigger and a Tyr was constructed to show that radical transfer in RNR enzymes occurs through a defined pathway of conserved aromatic residues [120]. The photooxidation of Trp was sufficient to initiate nucleotide reduction suggesting that the radical transfer to Cys439 was through Y731 of R1, with the peptide Tyr acting as Y356 of R2. The reaction is coupled to the reduction of dioxygen to hydrogen peroxide. A redox-active Trp residue was created as a result of multistep ET on a *de novo* helix bundle peptide [121]. The electron donation of the Trp residue to the photo-oxidized pyrene effectively competed with the back ET between pyrene⁺⁺ and methyl viologen⁺. Both pyrene and methyl viologen moieties were bound to the peptide. Tommos et al used two variants of *de novo* designed proteins of $\alpha 3W$ and $\alpha 4W$ to create radical proteins. The redox properties of a single Tyr or Trp in the two *de novo* designed radical proteins were probed by differential pulse voltammetry [122, 123]. The redox potentials were significantly elevated compared to the free Tyr and Trp in solution and this was related to the specific π - interaction of the Tyr and Trp side chains with the neighboring residues.

4. ELECTRON TRANSFER

ET in proteins typically occurs between protein bound cofactors, an electron donor cofactor and an electron acceptor cofactor, which are separated by large distances (5 to 20 Å). In general non-adiabatic theory, the electron is considered to switch between two states, a process commonly called electron tunneling. The 'tunneling' term invokes an image of a particle disappearing from one side of a barrier and re-appearing on the other, rather than one of spending time inside the barrier on passing through. In a protein ET reaction, energy is placed in an orbital on the donor cofactor by some means (eg. thermal, photochemical) and then the electron transfers to an orbital on the acceptor cofactor. The tunneling electron is coupled to the nuclear vibrational modes of both the surrounding protein and solvent as well as vibrational modes on the donor and acceptor moieties them-

selves. The nuclear vibrational modes are collectively called the bath. If the donor and acceptor states could be isolated from the bath, the electron would coherently transfer back and forth between the donor and acceptor orbitals. In reality, the tunneling electron loses energy to the bath which accounts for the polarization of the surrounding protein and solvent environment in response to the field of the transferring electron as well as shifts of vibrational modes on the donor and acceptor cofactors. Because the bath contains a very large number of modes, it is highly improbable that any energy transferred to the bath will be transferred back to the tunneling electron. The end result is that the electron moves from being localized on the donor to being localized on the acceptor, with some energy lost to the bath.

In classical Marcus theory [124], the rate constant for a non-adiabatic ET reaction is given by

$$k_{et} = (2\pi/h) T_{DA}^2 FC, \quad (1)$$

where T_{DA} is the electronic coupling matrix element and FC is the nuclear Franck-Condon factor. In a semi-classical treatment of the nuclear motions,

$$FC = (4\pi\lambda k_B T)^{-1/2} \exp(-(\Delta G^0 + \lambda)^2 / 4\lambda k_B T) \quad (2)$$

Here ΔG^0 is the free energy change due to ET, λ is the reorganization energy, k_B is the Boltzmann constant and T is temperature. According to Eq. (2), for a given T_{DA} the ET rate will be maximal when $-\Delta G^0 = \lambda$, i.e. the energy loss necessary to localize the electron on the acceptor equals the energy the bath needs to reorganize in response to the transfer of the electron. When $-\Delta G^0 = \lambda$ the reaction is activationless. As ΔG^0 becomes more negative, $-\Delta G^0 > \lambda$, according to classical Marcus theory the ET rate will slow. When $-\Delta G^0 < \lambda$ the ET reaction is said to be in Marcus inverted region. However, it is important to realize that the inverted region dependence of k_{et} on ΔG^0 is entirely reliant on the classical limit of the ET rate law Eq. 1. If the effects of nuclear tunneling between the reactant and the excited vibrational levels of the products are included, then k_{et} becomes considerably less sensitive to ΔG^0 in the inverted region [125].

The criteria for the ET to be in the non-adiabatic regime is that there is a small overlap of the electron donating orbital and the electron accepting orbital. In the case of proteins, the donor and acceptor cofactors are separated by such a large distance that there is a negligibly small probability of the electron directly tunneling between the two orbitals. Instead, the tunneling of the electron is facilitated by a bridging medium which lowers the effective barrier height between the donor and acceptor orbital by providing a series of virtual orbitals that the electron can transiently populate. The virtual orbitals on the bridge give rise to the concept of superexchange [126] which is typically applicable to ET reactions in proteins. The intervening protein medium between the donor and acceptor cofactors is considered to be the bridge.

The role of the protein bridge in protein ET has been the subject of numerous investigations, with the question being whether fast pathways through the protein bridge have been selected in evolution. Dutton and co-workers [127, 128] presented evidence that the protein bridge appears to be a uniform structure where $T_{DA} \sim \exp(-\beta R)$ as predicted by tunneling through a uniform square barrier. Here R is the 'edge to edge' distance between the donor and acceptor moieties.

Based on a the statistical analysis of available protein structures for which ET rates were measured, and assigning an edge-to-edge distance between cofactors, a value of $\beta \sim 1.4 \text{ \AA}^{-1}$ was suggested [128]. The explanation for an observed lack of fast tunneling pathways through the protein bridge is that any favorable (or unfavorable) characteristics of the protein bridge between the donor and acceptor tend to average out for protein ETs over large distances [129].

Gray and co-workers [130] presented evidence that some structural selection to increase T_{DA} was possible, for example, protein secondary structures such as an α -helix. In an extensive number of studies dating back since the early 1980's, Gray and co-workers investigated protein ET reactions using ruthenium (Ru)-modified metalloproteins [131]. Cyt *c*, Mb, azurin and cyt *b*₅ were modified through the attachment of Ru complexes (usually penta-amine-ruthenium or bis(bipyridine)-ruthenium-(imidazole)). Changes in the point of attachment of the Ru-complexes were used to determine the distance dependence of ET as well as the effect of the intervening medium in coupling electron donors and acceptors. Approaches to modify the driving force ΔG^0 and the reorganization energy λ included modification of the Ru-complexes as well as through metal-substitution (e.g., Zn, Mg, Cd, Pt, Pd) of the heme. The results from the experiments were interpreted in terms of a model for ET as a pathway of electron wavefunctions through specified bonds, each bond being represented by its length, Λ , multiplied by a coupling electronic factor σ [132]. The rate constant for ET, k_{et} , in terms of this model was given by: $\log k_{et} = \text{constant} \times \sum (\Lambda_i \sigma_i)$ where *i* is an atom on one path of *n* atoms. The pathway of bonds is selected as the path that maximized the value of T_{DA} . Winkler and Gray 1997 stressed differences between α -helical and β -structures (different H-bond σ terms were proposed) to explain deviations of measured ET in blue copper azurin proteins and cyt *c* from a fit to a line with slope $\beta \sim 1.4 \text{ \AA}^{-1}$ [133].

4.1. Electron Transfer in Artificial Proteins

Synthetic peptides have proven to be a valuable tool to investigate the role of the protein bridge in biological ET. A *de novo*-designed metalloprotein containing ruthenium-bipyridine complex and a heme was constructed and used to study the laser-induced ET across an α -helix [134]. The modular assembly of different amphiphilic helices by chemo-selective coupling to a cyclic peptide template was used to construct an antiparallel four-helix bundle. A heme was bound in a binding pocket by two ligating His residues and a ruthenium-tris(bipyridine) complex was covalently bound to different positions at the hydrophilic side of one of the heme-binding helices. The transient optical absorptions of the synthetic proteins were followed at 430 and 420 nm, respectively, to detect the slow recombination reaction $\text{Ru}^{\text{III}}\text{-heme(Fe}^{\text{II}}) \rightarrow \text{Ru}^{\text{II}}\text{-heme(Fe}^{\text{III}})$. A single exponential fit in the range from 5 to 80 μs yielded rate constants between 0.02 and 0.07 μs for heme-Ru-metalloprotein variants, respectively. However in attempting to explain the ET kinetics, the authors suggested further work was necessary using a shorter linkage to the ruthenium complex, to reduce uncertainties about the possible conformations and distances involved in the process.

Similarly, flash photolysis and pulse radiolysis measurements on a designed metalloprotein in both folded and unfolded states were used to investigate effects of helical conformation on ET rates [135]. A 16-mer helical bundle (three-helix metalloprotein) was modified with a capping $\text{Co}^{\text{III}}(\text{bipyridine})_3$ electron acceptor at the N-terminus and a 1-ethyl-1'-ethyl-4,4'-bipyridinium donor at the C-terminus. In the presence of 6 M urea, the random coil bundle (0% helicity) had an observed ET rate constant of $k_{et} = 900 \pm 100 \text{ s}^{-1}$ while in the presence of 25% trifluoroethanol, the helicity of the peptide was 80% and the k_{et} increased to $2000 \pm 200 \text{ s}^{-1}$. The increase in the rate constant was ascribed to the change in donor-acceptor distance between the random coil and helical bundle states. As a mimic for *in vivo* ET, a dimeric 31-mer synthetic metalloprotein containing two ruthenium complexes was designed to accomplish ET across the non-covalent peptide-peptide interface [136]. This design allowed Ogawa and co-workers to show a long-range intra-protein ET ($\sim 24 \text{ \AA}$) with a rate constant of 380 s^{-1} across a non-covalent peptide-peptide interface [137]. The rates of photoinduced ET from a pyrenyl group to a nitrophenyl group attached to an α -helical polypeptide were determined as function of the number of spacer amino acids and shown to be dependent on the edge-to-edge distance between the two chromophores [138]. It was also shown that a light-induced triplet ET between a Zn(II)-protoporphyrin bound to a His in the interior of a synthetic four-helix bundle protein and an anthraquinone molecule as a external electron acceptor occurred [139]. The protein was constructed by coupling four amphiphilic synthetic helical peptides via a linker to a cyclic decapeptide template.

In the α -helical structure, the dipole moments of the residues align in the same direction, nearly parallel to the helix axis. The α -helical structure can potentially be stabilized by changes in cofactor charges that give rise to dipole-dipole coupling. Also, proteins are known to fold or change conformation in response to a change in the charge distribution of their cofactors. Measurements by Lockhart and Kim [140, 141, 142] of the interaction between monomeric α -helices and solvent-exposed dipolar groups or titratable groups showed that an interaction energy of $>1 \text{ kcal mol}^{-1}$ existed between probe and partial charge on the helix N-terminus and was stabilizing to the α -helices. Tris-bipyridyl derivatives of Ru(II) were used by Huang et al to investigate the folding kinetics of a series of peptides and their Ru-modified derivatives [143]. The incorporation of an aromatic capping group led to a slight enhancement of the helical content and an enhanced stability of $\sim 0.15 \text{ kcal/mol}$ through the mechanism of dipole-dipole coupling. The permanent field arising from the helical dipole could have an effect on rapid ET between redox centers [144].

A directional ET was reported between *N,N*-dimethylaniline donor and pyrene acceptor attached to two peptides which differed in the positions of the donor and acceptor with respect to the helix ends [145]. The directional sensitivity was ascribed to the helix dipole moment; however, such an effect was not observed when ET between polar metal complexes in a metalloprotein was studied [146]. The two metal complexes were Ru polypyridyl donor and pentammine Ru (III) acceptor covalently attached to a 30-residue polypeptide via a Cys and a His, respectively. The

ET direction was reversed by switching the position of the Cys and His residues on the peptide. By varying the distance between two ruthenium ions attached to a three-helix bundle motif, a strong distance dependence for the ET rate was observed indicating the structural integrity of the designed protein [147]. It was also shown that ET occurred between a covalently attached flavin and a bis-His ligated heme in a light-activatable molecular maquette under continuous illumination [148].

4.2. Electron Transfer in Reaction Centers

Reaction centers of photosystems contain multiple redox cofactors with tuned redox potentials. The distances and orientations of cofactors are optimized such that a directional ET can be established while avoiding charge recombination. The use of multiple cofactors separated by $< 14\text{--}15 \text{ \AA}$ ensures that electron tunneling is the much more efficient mechanism of transfer than catalytic turnover [149]. The result is a charge-separated state with the oxidized and reduced cofactors well separated from each other. In the photosystems, the initial ET steps between primary electron donors and acceptors are extremely fast but these gradually slow as the probability of charge recombination becomes smaller. The oxidation potential of P is dependent on the type of photosystem and varies from 0.35 V in Chl *d*-containing PSI of *A. marina* [150] to $> 1.2 \text{ V}$ in PSII [151] to serve the particular photochemical function. In the purple bacterial RC, P₈₇₀, is a special pair of BChl *a* molecules with an E_m of +0.45 for the P₈₇₀⁺/P₈₇₀ couple. The BChl molecules in this red-shifted dimer are separated by $\sim 3 \text{ \AA}$ in the pyrrole ring I and by $\sim 7 \text{ \AA}$ from the ring centers. In PSII, the Chl *a* molecules in the P₆₈₀ dimer have a more monomeric nature and the ring centers are positioned about 8.5 \AA apart, generating the extremely strong oxidant ($E_m \sim + 1.2 \text{ V}$) needed to oxidize water to O₂ (+ 870 mV at pH 6). As a demonstration of additional mechanisms for tuning the redox potential, it was shown that the formation of up to four hydrogen bonds with the chlorin ring of the BChl special pair in a genetically modified *Rhodobacter sphaeroides* could additively increase the oxidation potential of P by a remarkable 355 mV [152, 153]. The mutated bacterial RC was then able to mimic the PSII donor side by oxidizing a nearby engineered Tyr residue [154] and a tightly bound Mn ion in a designed binding site [155]. Interestingly, the increase in the midpoint potential was found to be dependent on the chemical nature of the H-bond forming residue, with His being the most effective.

The quantum yield of charge separation is estimated to approach unity in the RCs of photosystems. The near perfect quantum yield that is achieved is due to the favorable ratios of the charge separation over the charge recombination reactions. Each ET step is at least 100 times faster than the associated charge recombination reaction, or decay of the initial excited state to the ground state by fluorescence or internal conversion. An explanation for the favorable ratios of the charge separation and recombination reactions is that the electronic coupling matrix element T_{DA} is large for the charge separation reaction and small for the recombination reaction. One way that RCs optimize the ratios of T_{DA} for charge separation and recombination is to break the charge separation process up into a series of small steps by using a set of electron carriers with extended π molecular orbitals.

The initial ET step, which must compete with the fastest recombination reaction, occurs through the accessory chlorophyll bridge such that the cofactors are essentially in contact. For charge separation to out-compete the recombination reaction, the product of the electronic coupling T_{DA} for each of the forward ET steps must be larger than the electronic coupling T_{DA} for the charge recombination ET step. Indirect recombination reactions that might return the system to an earlier radical pair are defeated by the loss of free energy at each step.

4.2.1. Artificial Reaction Centers

The design of a protein-based artificial RC capable of executing a stable charge separation must adhere to the same principle as natural RCs. The simplest design is a two-step ET where the first ET occurs from an excited photocatalyst to a properly positioned electron acceptor. The photocatalyst cation then accepts an electron from a properly positioned electron donor. If the electronic coupling T_{DA} between the electron donor and electron acceptor is smaller than the product of the electronic couplings between the photocatalyst and the electron acceptor/donor, then a stable charge separation can potentially be created. The design also needs to optimize the respective redox potentials to ensure indirect recombination reactions are defeated as in the natural systems. The electron acceptor can be a heme, a quinone or a Fe-S cluster while the electron donor can be a heme, a metal or a redox-active residue. The first example of a protein-based artificial RC was reported in Dutton's group [156]. The design was based on covalent attachment of two coproporphyrins at the two N-termini of peptides forming a four α -helix bundle forming a pair. The two coproporphyrins appear to be positioned on the exterior of the helix bundle allowing four bis-His ligated hemes to occupy the interior. The four α -helix bundle protein is a self-assembled dimer of helix-loop-helix structures each formed by two peptides (CGGG ELWKLHE ELLKKFE ELLKLHE RRLKKL-CoNH₂) linked through a disulfide bond in the loop region. To make these peptides photoactive, either Zn(II) derivative porphyrins were used or alternatively approach a Ru(II) porphyrin was incorporated into one of the heme binding pocket via His ligation. In the first design, the photoactive moiety is solvent exposed while in the second case it is placed in the interior. The effect of the distance between heme and photoactive electron donor was analysed in a series of designs and it was concluded that successful creation of charge separation would require a second redox cofactor, such as quinone, be incorporated in the system. In a similar approach, Cristian *et al.* reported a designed metalloprotein (aRC) that consists of a four-helix bundle functionalised with two bis-His-bound metal cofactors: a Ru(bpy)₂ moiety and a heme group. The tetrahelical scaffold is comprised of two identical 54-residue helix-loop-helix peptides joined through a disulfide bond between the C-terminal cysteines to yield an antiparallel four-helix bundle. Photoexcitation of this aRC led to efficient quenching of the luminescence of the ruthenium complex by the heme group. The ET rate was estimated from steady-state luminescence measurements to be of the order $5 \times 10^{10} \text{ s}^{-1}$. The charge recombination reaction was three orders of magnitude slower, enabling the aRC to oxidize cyt *c* and reduce naphthoquinone-2-sulfonate in solution.

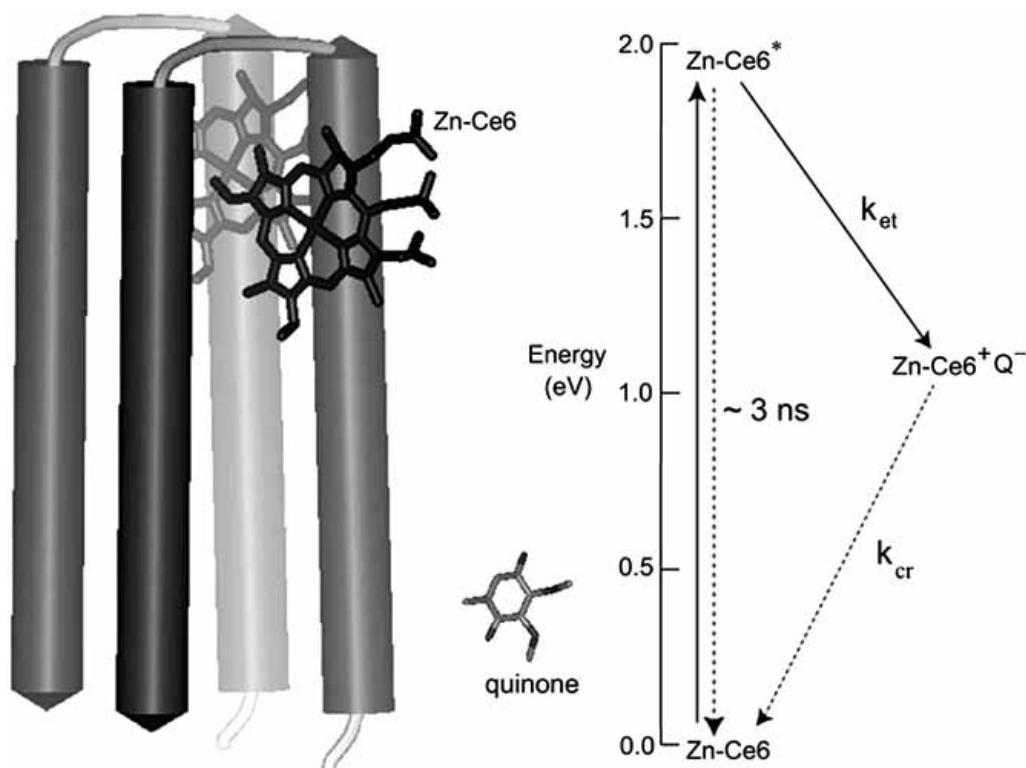


Fig. (4). A cartoon presentation of an artificial protein-based RC (A) and the ET process (B). The artificial RC can be either a water-soluble protein carrying a photocatalyst such as a chlorin or a membrane-bound protein carrying a Chl molecule. Quinone can be used as the electron acceptor either covalently-bound to the protein or free like mobile plastoquinones. Upon light excitation of the photocatalyst, an electron is transferred to the quinone with a rate constant of k_{et} . In the absence of a third cofactor, functioning as the electron donor to the oxidized photocatalyst, a charge recombination occurs with a rate constant of k_{cr} . The charge recombination rate that controls the lifetime of the initial charge-separated state is dependent on the redox potentials and distance between cofactors and may be manipulated as a consequence.

The design of an artificial RC can take advantage of Chl derivatives that can be bound to proteins much more simply than Chl (Fig. 4). The Chl derivatives can be made by enzymatic reactions or by using organic chemistry methods to remove the phytol tail while maintaining the chlorin ring. Some examples of these compounds are chlorin *e6* (*Ce6*), pyropheophorbide and tetracarboxyphenylchlorin. These compounds, like their parent molecule of Chl, donate or accept electrons from extensive conjugated π molecular orbitals and have small reorganization energies because the uptake or release of an electron can be accommodated with only minimal structural changes. They are photoactive but with more desirable physicochemical properties compared to Chls. The more desirable physicochemical properties include being more stable than Chls and, most importantly, aggregation can be controlled allowing for specific binding to engineered binding sites in a protein. In these cases, binding modes other than His ligation become feasible. For example, using a chemo-selective method, a Zn-pheophorbide *b* derivative was covalently attached to a synthetic four-helix bundle protein through an aldehyde group and a modified lysine residue [157]. Similarly, it was also shown that light-induced ET could occur between a chromophore bound in the interior of a synthetic four-helix bundle protein and exogenously added anthraquinone [139]. We reported a His-dependent binding of Zn-*Ce6* to a synthetic four-helix bundle protein (Fig. 4) [158]. By preparing two variants of the peptide lacking either one or both of the His residues, we

showed that each monomer had two bound Zn-*Ce6* molecules. Light-induced ET between Zn-*Ce6* and a quinone as the external electron acceptor was established and measured using time-resolved EPR spectroscopy monitoring the oxidation of Zn-*Ce6* and its reduction with the quinone. The kinetics showed a bi-phasic behaviour with a very fast component lasting through the flash duration. More studies were performed to characterise the photochemistry of the system using various quinones and chlorin derivatives [159]. The very fast kinetic component is now attributed to the ET between Zn-*Ce6* and quinones that entered into the interior of the complex. A similar strategy was used to modify Cyt b_{562} by replacing the heme with the Zn-*Ce6* [160]. In this design, the quinone was covalently bound to the interior of the protein through an engineered Cys residue. Fluorescence studies showed ET between the Zn-*Ce6* and the bound quinone.

CONCLUSION

In this review, we have summarized aspects of functional design for artificial proteins, particularly in relation to receptor binding, metal and porphyrin binding, and construction of photoactive proteins. The literature in this area is already large and rapidly increasing and this review surveys the more recent developments. Future challenges for this exciting field of protein biochemistry are likely to include a substantial expansion in the range of design strategies used to mimic ever wider classes of natural proteins, but also the creation of more complex *de novo* artificial proteins with the ability to

bind multiple cofactors. The design of small proteins containing multiple cofactors with native-like structures will require the development of more sophisticated design principles than those being employed currently. Such proteins would be able to mimic some functions of large membrane-bound oxidoreductases, which are capable of multi-step ET reactions.

ACKNOWLEDGEMENTS

This work was supported by Australian Research Council Discovery Grant (DP0450421). BBW acknowledges the support of an Australian Postgraduate Award.

ABBREVIATIONS

(B)Chl	=	(Bacterio)chlorophyll
(B)Pheo	=	(Bacterio)pheophytin
AP3	=	An amphiphilic protein maquette
BBA	=	A homotetrameric miniprotein
Bc-Csp	=	A thermophilic cold shock protein
Bs-CspB	=	A mesophilic cold shock protein
C34	=	The C-terminal region of HIV-1 gp41
CCSL	=	A coiled coil miniprotein
Ce6	=	Chlorin e6
DF1	=	Due-ferri-1 protein
DPP-Fe(III)	=	Fe(III)-diphenylporphyrin
E_m	=	Midpoint potential
ET	=	Electron transfer
Fe-S	=	Iron-sulfur
IL1	=	Interleukin-1
IL4	=	Interleukin-4
IL4-R α	=	IL4 α receptor
IL5	=	Interleukin-5
LH	=	Light harvesting
LH1	=	Bacterial core LH complex
LP	=	Lipophilic
Mb	=	Myoglobin
P	=	Primary electron donor
PCP	=	Peridinin-chlorophyll <i>a</i> -protein
PSI	=	Photosystem I
PSII	=	Photosystem II
RC	=	Reaction centre
WSCP	=	Water-soluble Chl binding protein.

REFERENCES

- Roder, H. and Colon, W. (1997) *Curr. Opin. Struct. Biol.*, 7, 15-28.
- Jackson, S.E. (1998) *Fold. Des.*, 3, R 81-R 91.
- Lindorff-Larsen, K., Rogen, P., Paci, E., Vendruscolo, M. and Dobson, C.M. (2005) *Trends Bioch. Sci.*, 30, 13-19.
- Feng, H.Q., Zhou, Z. and Bai, Y.W. (2005) *Proc. Natl. Acad. Sci. USA*, 102, 5026-5031.
- Tsai, C.J., Ma, B.Y., Kumar, S., Wolfson, H. and Nussinov, R. (2001) *Crit. Rev. Biochem. Mol. Biol.*, 36, 399-433.
- Korkegian, A., Black, M.E., Baker, D. and Stoddard, B.L. (2005) *Science*, 308, 857-860.
- Perl, D., Mueller, U., Heinemann, U. and Schmid, F.X. (2000) *Nat. Struct. Biol.*, 7, 380-383.
- House-Pompeo, K., Xu, Y., Joh, D., Speziale, P. and Höök, M. (1996) *J. Biol. Chem.*, 271, 1379-1384.
- Schweers, O., Schönbunn-Hanebeck, E., Marx, A. and Mandelkow, E. (1994) *J. Biol. Chem.*, 269, 24290-24297.
- Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H.R., Müller-Frohne, M., Zirwer, D., Czarbecki, J. and Damaschun, G. (1995) *Biochemistry*, 34, 13211-13218.
- Kriwacki, R.W., Hengst, L., Tennant, L., Reed, S.I. and Wright, P.E. (1996) *Proc. Natl. Acad. Sci. USA*, 93, 11504-9.
- Mayo, K.H. (2000) *Trends Biotech.*, 18, 212-217.
- Hill, R.B., Raleigh, D.P., Lombardi, A. and DeGrado, W.F. (2000) *Acc. Chem. Res.*, 33, 745-54.
- Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. and Hecht, M.H. (1993) *Science*, 262, 1680-1685.
- Schafmeister, C.E., Miercke, L.J. and Stroud, R.M. (1993) *Science*, 262, 734-8.
- Weber, P.C. and Salemme, F.R. *Nature*, 287, 82-84.
- Gibney, B.R., Rabanal, F., Skalicky, J.J., Wand, A.J. and Dutton, L.P. (1999) *J. Am. Chem. Soc.*, 121, 4952-4960.
- Hill, R.B. and DeGrado, W. (1998) *J. Am. Chem. Soc.*, 120, 1138-1145.
- Skalicky, J.J., Gibney, B.R., Rabanal, F., Bieber-Urbauer, R.J., Dutton, P.L. and Wand, A.J. (1999) *J. Am. Chem. Soc.*, 121, 4941-4951.
- Walsh, S.T.R., Cheng, H., Bryson, J.W., Roder, H. and DeGrado, W.F. (1999) *Proc. Natl. Acad. Sci. USA*, 96, 5486-5491.
- Schafmeister, C.E., LaPorte, S.L., Miercke, L.J. and Stroud, R.M. (1997) *Nature Struct. Biol.*, 4, 1039-46.
- Stotz, C.E. and Topp, E.M. (2004) *J. Pharm. Sci.*, 93, 2881-2894.
- Kortemme, T., Ramirezalvarado, M. and Serrano, L. (1998) *Science*, 281, 253-256.
- Griffiths-Jones, S.R. and Seale, M.S. (2000) *J. Am. Chem. Soc.*, 122, 8350-8356.
- Ali, M.H., Peisach, E., Allen, K.N. and Imperiali, B. (2004) *Proc. Natl. Acad. Sci. USA*, 101, 12183-12188.
- Ali, M.H., Taylor, C.M., Grigoryan, G., Allen, K.N., Imperiali, B. and Keating, A.E. (2005) *Structure*, 13, 225-234.
- Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L. and Baker, D. (2003) *Science*, 302, 1364-1368.
- Domingues, H., Cregut, D., Sebald, W., Oschkinat, H. and Serrano, L. (1999) *Nature Struct. Biol.*, 6, 652-656.
- LaPorte, S.L., Forsyth, C.M., Cunningham, B.C., Miercke, L.J., Akhavan, D. and Stroud, R.M. (2005) *Proc. Natl. Acad. Sci. USA*, 102, 1889-1894.
- Sia, S.K. and Kim, P.S. (2003) *Proc. Natl. Acad. Sci. USA*, 100, 9756-9761.
- Scherf, T., Kasher, R., Balass, M., Fridkin, M., Fuchs, S. and Katchalski-Katzir, E. (2001) *Proc. Natl. Acad. Sci. USA*, 98, 6629-6634.
- Li, C.Z., Plugariu, C.G., Bajgier, J., White, J.R., Liefer, K.M., Wu, S.J. and Chaiken, I. (2002) *J. Mol. Recognit.*, 15, 33-43.
- Milburn, M.V., Hassell, A.M., Lambert, M.H., Jordan, S.R., Proudfoot, A.E., Graber, P. and Wells, T.N. (1993) *Nature*, 363, 172-6.
- Myszka, D.G. and Chaiken, I.M. (1994) *Biochemistry*, 33, 2363-2372.
- Yanofsky, S.D., Baldwin, D.N., Butler, J.H., Holden, F.R., Jacobs, J.W., Balasubramanian, P., Chinn, J.P., Cwirla, S.E., Peters-Bhatt, E., Whitehorn, E.A., Tate, E.H., Akeson, A., Bowlin, T.L., Dower, W.J. and Barrett, R.W. (1996) *Proc. Natl. Acad. Sci. USA*, 93, 7381-6.
- Feng, J.N., Li, Y. and Shen, B.F. (2004) *Peptides*, 25, 1123-1131.
- Martin, L., Stricher, F., Misse, D., Sironi, F., Pugnieri, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., Menez, A., Lusso, P., Veas, F. and Vita, C. (2003) *Nature Biotech.*, 21, 71-76.
- Huang, C.C., Stricher, F., Martin, L., Decker, J.M., Majeed, S., Barthe, P., Hendrickson, W.A., Robinson, J., Roumestand, C., Sodroski, J., Wyatt, R., Shaw, G.M., Vita, C. and Kwong, P.D. (2005) *Structure*, 13, 755-768.

- [39] Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., Roumestand, C., Yang, Y.S., Ylisastigui, L., Benjouad, A. and Gluckman, J.C. (1999) *Proc. Natl. Acad. Sci. USA*, *96*, 13091-13096.
- [40] Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (1998) *Nature*, *393*, 648-659.
- [41] Martin, L., Barthe, P., Combes, O., Roumestand, C. and Vita, C. (2000) *Tetrahedron*, *56*, 9451-9460.
- [42] Holm, R.H., Kennepohl, P. and Solomon, E.I. (1996) *Chem. Rev.*, *96*, 2239-2314.
- [43] Degtyarenko, K. (2000) *Bioinformatics*, *16*, 851-864.
- [44] Wan, L., Twitchett, M.B., Eltis, L.D., Mauk, A.G. and Smith, M. (1998) *Proc. Natl. Acad. Sci. USA*, *95*, 12825-31.
- [45] Arisawa, A., Lin, Z., Arnold, F.H. and Joo, H. (1999) *Chem. Biol.*, *6*, 699-706.
- [46] Shifman, J.M., Gibney, B.R., Sharp, R.E. and Dutton, P.L. (2000) *Biochemistry*, *39*, 14813-14821.
- [47] Blanchard, L., Marion, D., Pollock, B., Voordouw, G., Wall, J., Bruschi, M. and Guerlesquin, F. (1993) *Eur. J. Biochem.*, *109*, 10912-10922.
- [48] Blackledge, M.J., Guerlesquin, F. and Marion, D. (1996) *Proteins*, *24*, 178-94.
- [49] Hay, S. and Wydrzynski, T. (2005) *Biochemistry*, *44*, 431-439.
- [50] Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) *Nature*, *426*, 413-418.
- [51] Moffet, D.A., Case, M.A., House, J.C., Voge, K., Williams, R.D., Spiro, T.G., McLendon, G.L. and Hecht, M.H. (2001) *J. Am. Chem. Soc.*, *123*, 2109-2115.
- [52] Moffet, D.A., Certain, L.K., Smith, A.J., Kessel, A.J., Beckwith, A. and Hecht, M.H. (2000) *J. Am. Chem. Soc.*, *122*, 7612-7613.
- [53] Chu, R., Takei, J., Knowlton, J.R., Andrykovitch, M., Pei, W., Kajava, A.V., Steinbach, P.J., Ji, X. and Bai, Y. (2002) *J. Mol. Biol.*, *323*, 253-62.
- [54] Garcia, P., Bruix, M., Rico, M., Ciofi-Baffoni, S., Banci, L., Shastri, M.C.R., Roder, H., Woodyear, T.D., Johnson, C.M., Fersht, A.R. and Barker, P.D. (2005) *J. Mol. Biol.*, *346*, 331-344.
- [55] Faraone-Mennella, J., Gray, H.B. and Winkler, J.R. (2005) *Proc. Natl. Acad. Sci. USA*, *102*, 6315-6319.
- [56] Cochran, F.V., Wu, S.P., Wang, W., Nanda, V., Saven, J.G., Therien, M.J. and DeGrado, W.F. (2005) *J. Am. Chem. Soc.*, *127*, 1346-1347.
- [57] Huang, S.S., Koder, R.L., Lewis, M., Wand, A.J. and Dutton, P.L. (2004) *Proc. Natl. Acad. Sci. USA*, *101*, 5536-41.
- [58] Ghirlanda, G., Osyczka, A., Liu, W., Antolovich, M., Smith, K.M., Dutton, P.L., Wand, A.J. and DeGrado, W.F. (2004) *J. Am. Chem. Soc.*, *126*, 8141-7.
- [59] Hillier, W. and Babcock, G.T. (2001) *Plant Physiol.*, *125*, 33-37.
- [60] Biesiadka, J., Loll, B., Kern, J., Irrgang, K.D. and Zouni, A. (2004) *Phys. Chem. Chem. Phys.*, *6*, 4733-4736.
- [61] Deisenhofer, J., Epp, O., Sinning, I. and Michel, H. (1995) *J. Mol. Biol.*, *246*, 429-457.
- [62] Ben-Shem, A., Frolov, F. and Nelson, N. (2003) *Nature*, *426*.
- [63] Kamiya, N. and Shen, J.-R. (2003) *Proc. Natl. Acad. Sci. USA*, *100*, 98-103.
- [64] Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J. and Iwata, S. (2004) *Science*, *303*, 1831-1838.
- [65] Zouni, A., Witt, H.-T., Kern, J., Fromme, P., Krauß, N., Saenger, W. and Orth, P. (2001) *Nature*, *409*, 739-743.
- [66] Liu, Z., H., Y., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X. and Chang, W. (2004) *Nature*, *428*, 287-292.
- [67] Kuhlbrandt, W., Wang, D. and Fujiyoshi (1994) *Nature*, *367*, 614-621.
- [68] Chirino, A.J., Lous, E.J., Huber, M., Allen, J.P., Schenck, C.C., Paddock, M.L., Feher, G. and Rees, D.C. (1994) *Biochemistry*, *33*, 4584-4593.
- [69] Xiong, L., Seibert, M., Gusev, A.V., Wasielewski, M.R., Hemann, C., Hille, C.R. and Sayre, R.T. (2004) *J. Phys. Chem. B*, *108*, 16904-16911.
- [70] Bassi, R., Croce, R., Cugini, D. and Sandona, D. (1999) *Proc. Natl. Acad. Sci. USA*, *96*, 10056-10061.
- [71] Satoh, H., Uchida, A., Nakayama, K. and Okada, M. (2001) *Plant Cell Physiol.*, *42*, 906-911.
- [72] Hofmann, E., P., W., Sharples, F.P., Hiller, R.G., Welte, W. and Diederichs, K. (1996) *Science*, *272*, 1788-1791.
- [73] Schmidt, K., Fufezan, C., Krieger-Liszskay, A., Satoh, H. and Paulsen, H. (2003) *Biochemistry*, *42*, 7427-7433.
- [74] Kashiwada, A., Nishino, N., Wang, Z.-Y., Nozawa, T., Kobayashi, M. and Nango, M. (1999) *Chem. Lett.*, *2*, 1301-1302.
- [75] Dudkowiak, A., Kusumi, T., Nakamura, C. and Miyake, J. (1999) *J. Photochem. Photobiol. B. Biol.*, *129*, 51-55.
- [76] Meadows, K.A., Parkes-Loach, P.S., Kehoe, J.W. and Loach, P.A. (1998) *Biochemistry*, *37*, 3411-3417.
- [77] Kehoe, J.W., Meadows, K.A., Parkesloach, P.S. and Loach, P.A. (1998) *Biochemistry*, *37*, 3418-3428.
- [78] Eggink, L.L. and Hooper, J.K. (2000) *J. Biol. Chem.*, *275*, 9087-9090.
- [79] Discher, B.M., Noy, D., Strzalka, J., Ye, S.X., Moser, C.C., Lear, J.D., Blasie, J.K. and Dutton, P.L. (2005) *Biochemistry*, *44*, 12329-12343.
- [80] Noy, D., Discher, B.M., Rubtsov, I.V., Hochstrasser, R.M. and Dutton, P.L. (2005) *Biochemistry*, *44*, 12344-12354.
- [81] Karlin, S., Zhu, Z.-Y. and Karlin, K.D. (1997) *Proc. Natl. Acad. Sci. USA*, *94*, 14225-14230.
- [82] Gray, H.B., Malmstrom, B.G. and Williams, R.J.P. (2000) *J. Biol. Inorg. Chem.*, *5*, 551-559.
- [83] Torrado, A., Walkup, G.K. and Imperiali, B. (1998) *J. Am. Chem. Soc.*, *120*, 609-610.
- [84] Ghadiri, M.R. and Choi, C. (1990) *J. Am. Chem. Soc.*, *112*, 1632-1634.
- [85] Daugherty, R.G., Wasowicz, T., Gibney, B.R. and DeRose, V.J. (2002) *Inorg. Chem.*, *41*, 2623-2632.
- [86] Xue, Y., Okvist, M., Hansson, O. and Young, S. (1998) *Protein Sci.*, *7*, 2099-2105.
- [87] Calhoun, J.R., Nastri, F., Maglio, O., Pavone, V., Lombardi, A. and DeGrado, W.F. (2005) *Biopolymers*, *80*, 264-78.
- [88] Stephens, P.J., Jollie, D.R. and Warshel, A. (1996) *Chem. Rev.*, *96*, 2491-2513.
- [89] Gibney, B.R., Mulholland, S.E., Rabanal, F. and Dutton, P.L. (1996) *Proc. Natl. Acad. Sci. USA*, *93*, 15041-15046.
- [90] Mulholland, S.E., Gibney, B.R., Rabanal, F. and Dutton, P.L. (1998) *J. Am. Chem. Soc.*, *120*, 10296-10302.
- [91] Mulholland, S.E., Gibney, B.R., Rabanal, F. and Dutton, P.L. (1999) *Biochemistry*, *38*, 10442-10448.
- [92] Regan, L. and DeGrado, W.F. (1988) *Science*, *241*, 976-978.
- [93] Scott, M.P. and Biggins, J. (1997) *Protein Sci.*, *6*, 340-346.
- [94] Laplaza, C.E. and Holm, R.H. (2001) *J. Am. Chem. Soc.*, *123*, 10255-10264.
- [95] Sow, T.C., Pedersen, M.V., Christensen, H.E.M. and Ooi, B.L. (1996) *Biochem. Biophys. Res. Com.*, *223*, 360-364.
- [96] Kharenko, O.A., Kennedy, D.C., Demeler, B., Maroney, M.J. and Ogawa, M.Y. (2005) *J. Am. Chem. Soc.*, *127*, 7678-7679.
- [97] Dismukes, G.C. (1996) *Chem. Rev.*, *96*, 2909-2926.
- [98] Zamocky, M. and Koller, F. (1999) *Prog. Biophys. Mol. Biol.*, *72*, 19-66.
- [99] Pace, R.J., in *Artificial Photosynthesis: From Basic Biology to Industrial Application* Collings, A. F., Critchley, C., Eds. (WILEY-VCH2, Weinheim, 2005).
- [100] Debus, R.J. (2001) *Biochim. Biophys. Acta*, *1503*, 164-186.
- [101] Pace, R.J. and Ahrling, K.A. (2004) *Biochim. Biophys. Acta*, *1655*, 172-178.
- [102] Costanzo, L.D., Wade, H., Geremia, S., Randaccio, L., Pavone, V., DeGrado, W.F. and Lombardi, A. (2001) *J. Am. Chem. Soc.*, *123*, 12749-12757.
- [103] DeGrado, W.F., Costanzo, L.D., Geremia, S., Lombardi, A., Pavone, V. and Randaccio, L. (2003) *Angew. Chem. Int. Ed.*, *42*, 417-420.
- [104] Kaplan, J. and DeGrado, W.F. (2004) *Proc. Natl. Acad. Sci. USA*, *101*, 11566-70.
- [105] Metz, J.G., Nixon, P.J., Rögner, M., Brudvig, G.W. and Diner, B.A. (1989) *Biochemistry*, *28*, 6960-6969.
- [106] Debus, R.J., Barry, B.A., Sithole, I., Babcock, G.T. and McIntosh, L. (1988) *Biochemistry*, *27*, 9071-9074.
- [107] Hoganson, C.W. and Babcock, G.T. (1997) *Science*, *277*, 1953-6.
- [108] Razeghifard, M.R., Chen, M., Hughes, J.L., Freeman, J., Krausz, E. and Wydrzynski, T. (2005) *Biochemistry*, *44*, 11178-11187.
- [109] Razeghifard, M.R., Klughammer, C. and Pace, R.J. (1997) *Biochemistry*, *36*, 86-92.
- [110] Razeghifard, M.R. and Pace, R.J. (1997) *Biochim. Biophys. Acta*, *1322*, 141-150.
- [111] Razeghifard, M.R., Wydrzynski, T., Pace, R.J. and Burnap, R.L. (1997) *Biochemistry*, *36*, 14474-14478.

- [112] Razeghifard, M.R. and Pace, R.J. (1999) *Biochemistry*, 38, 1252-1257.
- [113] Yoshikawa, S., Shinzawaitoh, K., Nakashima, R., Yaono, R., Yamashita, E., Inoue, N., Yao, M., Fei, M.J., Libeu, C.P., Mizushima, T., Yamaguchi, H., Tomizaki, T. and Tsukihara, T. (1998) *Science*, 280, 1723-1729.
- [114] Proshlyakov, D.A., Pressler, M.A., DeMaso, C., Leykam, J.F., DeWitt, D.L. and Babcock, G.T. (2000) *Science*, 290, 1588-1591.
- [115] Whittaker, J.W. (2005) *Arch. Biochem. Biophys.*, 433, 227-239.
- [116] Högbom, M., Stenmark, P., Voevodskaya, N., McClarty, G., Graslund, A. and Nordlund, P. (2004) *Science*, 305, 245-248.
- [117] Page, C.C., Moser, C.C. and Dutton, P.L. (2003) *Curr. Opin. Chem. Biol.*, 7, 551.
- [118] Gray, H.B. and Winkler, J.R. (2003) *Q. Rev. Biophys.*, 36, 341.
- [119] Aubert, C., Vos, M.H., Mathis, P., Eker, A.P.M. and Brettel, K. (2000) *Nature*, 405, 586.
- [120] Chang, M.C.Y., Yee, C.S., Stubbe, J. and Nocera, D.G. (2004) *Proc. Natl. Acad. Sci. USA*, 101, 6882-6887.
- [121] Jones, G., Vullev, V., Braswell, E.H. and Zhu, D. (2000) *J. Am. Chem. Soc.*, 122, 388-389.
- [122] Tommos, C., Skalicky, J.J., Pilloud, D.L., Wand, A.J. and Dutton, P.L. (1999) *Biochemistry*, 38, 9495-9507.
- [123] Dai, Q.-H., Tommos, C., Fuentes, E.J., Blomberg, M.R.A., Dutton, P.L. and Wand, A.J. (2002) *J. Am. Chem. Soc.*, 124, 10952-10953.
- [124] Marcus, R.A. and Sutin, N. (1985) *Biochim. Biophys. Acta*, 811, 265-322.
- [125] Jortner, J. (1979) *J. Chem. Phys.*, 64, 4860-4871.
- [126] McConnell, H. (1961) *J. Chem. Phys.*, 35, 508-515.
- [127] Farid, R.S., Moser, C.C. and Dutton, P.L. (1993) *Curr. Opin. Struct. Biol.*, 3, 225-233.
- [128] Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S. and Dutton, P.L. (1992) *Nature*, 355, 796-802.
- [129] Larsson, S. (1998) *Biochim. Biophys. Acta*, 1365, 294-300.
- [130] Beratan, D.S., Onuchic, J.N., Winkler, J.R. and Gray, H.B. (1992) *Science*, 258.
- [131] Gray, H.B. and Winkler, J.R. (2005) *Proc. Natl. Acad. Sci. USA*, 102, 3534-3539.
- [132] Beratan, D.N., Betts, J.N. and Onuchic, J.N. (1991) *Science*, 252, 1285-1288.
- [133] Winkler, J.R. and Gray, H.B. (1997) *J. Biol. Inorg. Chem.*, 2, 399-404.
- [134] Rau, H.K., DeJonge, N. and Haehnel, W. (1998) *Proc. Natl. Acad. Sci. USA*, 95, 11526-11531.
- [135] Mutz, M.W., McLendon, G.L., Wishart, J.F., Gaillard, E.R. and Corin, A.F. (1996) *Proc. Natl. Acad. Sci. USA*, 93, 9521-9526.
- [136] Kozlov, G.V. and Ogawa, M.Y. (1997) *J. Am. Chem. Soc.*, 119, 8377-8378.
- [137] Kornilova, A., Wishart, J., Xiao, W., Lasey, R., Fedorova, A., Shin, Y. and Ogawa, M. (2000) *J. Am. Chem. Soc.*, 122, 7999-8006.
- [138] Sisido, M., Hoshino, S., Kusano, H., Kuragaki, M., Makino, M., Sasaki, H., Smith, T.A. and Ghiggino, K.P. (2001) *J. Phys. Chem. B*, 105, 10407-10415.
- [139] Fahnenschmidt, M., Bittl, R., Schlodder, E., Haehnel, W. and Lubitz, W. (2001) *Phys. Chem. Chem. Phys.*, 3, 4082-4090.
- [140] Lockhart, D.J., Sharp, K.A. and Honig, B. (1994) *Biophys. J.*, 67, 2251-60.
- [141] Kim, P.S. and Lockhart, D.J. (1992) *Science*, 257, 947-51.
- [142] Kim, P.S. and Shoemaker, K.R. (1993) *Science*, 260, 198-202.
- [143] Huang, C.-Y., He, S., DeGrado, W.F., McCafferty, D.G. and Gai, F. (2002) *J. Am. Chem. Soc.*, 124, 12674-12675.
- [144] Alegria, G. and Dutton, P.L. (1991) *Biochim. Biophys. Acta*, 1057, 239-257.
- [145] Fox, M.A. and Galoppini, E. (1997) *J. Am. Chem. Soc.*, 119, 5277-5285.
- [146] Fedorova, A., Chaudhari, A. and Ogawa, M. (2003) *J. Am. Chem. Soc.*, 125, 357-362.
- [147] Mutz, M.W., Case, M.A., Wishart, J.F., Ghadiri, M.R. and McLendon, G.L. (1999) *J. Am. Chem. Soc.*, 121, 858-859.
- [148] Sharp, R.E., Moser, C.C., Rabanal, F. and Dutton, P.L. (1998) *Proc. Natl. Acad. Sci. USA*, 95, 10465-10470.
- [149] Page, C.C., Moser, C.C., Chen, X. and Dutton, P.L. (1999) *Nature*, 402, 47-52.
- [150] Hu, Q., Miyashita, H., Iwasaki, I., Kurano, N., Miyachi, S., Iwaki, M. and Itoh, S. (1998) *Proc. Natl. Acad. Sci. USA*, 95, 13319-13323.
- [151] Rappaport, F., Guergova-Kuras, M., Nixon, P.J., Diner, B.A. and Lavergne, J. (2002) *Biochemistry*, 41, 8518-27.
- [152] Lin, X., Murchison, H.A., Nagarajan, V., Parson, W.W., Allen, J.P. and Williams, J.C. (1994) *Proc. Natl. Acad. Sci. USA*, 91, 10265-10269.
- [153] Ivancich, A., Artz, K., Williams, J.C., Allen, J.P. and Mattioli, T.A. (1998) *Biochemistry*, 37, 11812-11820.
- [154] Kálmán, L., LoBrutto, R., Allen, J.P. and Williams, J.C. (1999) *Nature*, 402, 696-699.
- [155] Thielges, M., Uyeda, G., Cámara-Artigas, A., Kálmán, L., Williams, J.C. and Allen, J.P. (2005) *Biochemistry*, 44, 7389-7394.
- [156] Rabanal, F., Gibney, B.R., DeGrado, W.F., Moser, C.C. and Dutton, P.L. (1996) *Inorg. Chim. Acta*, 243, 213-218.
- [157] Rau, H.K., Snigula, H., Struck, A., Robert, B., Scheer, H. and Haehnel, W. (2001) *Eur. J. Biochem.*, 268, 3284-3295.
- [158] Razeghifard, M.R. and Wydrzynski, T.J. (2003) *Biochemistry*, 42, 1024-1030.
- [159] Razeghifard, R. and Wydrzynski, T. (2005) *In preparation*.
- [160] Hay, S., Wallace, B.B., Smith, T.A., Ghiggino, K.P. and Wydrzynski, T. (2004) *Proc. Natl. Acad. Sci. USA*, 101, 17675-17680.
- [161] Redfield, C., Smith, L.J., Boyd, J., Lawrence, G.M.P., Edwards, R.G., Gershater, C.J., Smith, R.A.G. and Dobson, C.M. (1994) *J. Mol. Biol.*, 238, 23-41.
- [162] Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1993) *Biochemistry*, 32, 6744-6762.
- [163] Razeghifard, M.R. (2004) *Protein Express. Purif.*, 37, 180-186.
- [164] Hage, T., Sebald, W. and Reinemer, P. (1999) *Cell*, 97, 271-281.
- [165] Hulsmeier, M., Scheufler, C. and Dreyer, M.K. (2001) *Acta Cryst.*, D57, 1334-1336.