Design principles for chlorophyll-binding sites in helical proteins

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · FEBRUARY 2011

Impact Factor: 2.63 · DOI: 10.1002/prot.22895 · Source: PubMed

8 AUTHORS, INCLUDING:

Paula Braun
Munich University of Applied Sciences
24 PUBLICATIONS **563** CITATIONS
SEE PROFILE

Vikas Nanda
Rutgers, The State University of New Jersey
72 PUBLICATIONS **1,712** CITATIONS
SEE PROFILE

Ronald Koder
City College of New York
46 PUBLICATIONS **773** CITATIONS
SEE PROFILE

Dror Noy
Migal - Galilee Technology Center
31 PUBLICATIONS **1,027** CITATIONS
SEE PROFILE

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Design principles for chlorophyll-binding sites in helical proteins

Paula Braun,[1] Eran Goldberg,[2] Christopher Negron,[3] Mathias von Jan,[4] Fei Xu,[5] Vikas Nanda,[5] Ronald L. Koder,[5] and Dror Noy[2]*

[1] Ludwig-Maximilians-Universität München, Department Biologie I, Botany, D-82152 Planegg-Martinsried, Germany
[2] Plant Sciences Department, Weizmann Institute of Science, Rehovot 76100, Israel
[3] Department of Physics, The City College of New York, New York, New York 10031
[4] DSMZ—German Collection of Microorganisms and Cell Cultures GmbH, D-38124 Braunschweig, Germany
[5] Robert Wood Johnson Medical School—UMDNJ Biochemistry, Center for Advanced Biotechnology and Medicine, Piscataway, New Jersey 08854

ABSTRACT

The cyclic tetrapyrroles, viz. chlorophylls (Chl), their bacterial analogs bacteriochlorophylls, and hemes are ubiquitous cofactors of biological catalysis that are involved in a multitude of reactions. One systematic approach for understanding how Nature achieves functional diversity with only this handful of cofactors is by designing *de novo* simple and robust protein scaffolds with heme and/or (bacterio)chlorophyll [(B)Chls]-binding sites. This strategy is currently mostly implemented for heme-binding proteins. To gain more insight into the factors that determine heme-/(B)Chl-binding selectivity, we explored the geometric parameters of (B)Chl-binding sites in a nonredundant subset of natural (B)Chl protein structures. Comparing our analysis to the study of a nonredundant database of heme-binding helical histidines by Negron *et al.* (Proteins 2009;74:400–416), we found a preference for the m-rotamer in (B)Chl-binding helical histidines, in contrast to the preferred t-rotamer in heme-binding helical histidines. This may be used for the design of specific heme- or (B)Chl-binding sites in water-soluble helical bundles, because the rotamer type defines the positioning of the bound cofactor with respect to the helix interface and thus the protein-binding site. Consensus sequences for (B)Chl binding were identified by combining a computational and database-derived approach and shown to be significantly different from the consensus sequences recommended by Negron *et al.* (Proteins 2009;74:400–416) for heme-binding helical proteins. The insights gained in this work on helix-(B)Chls-binding pockets provide useful guidelines for the construction of reasonable (B)Chl-binding protein templates that can be optimized by computational tools.

Proteins 2011; 79:463–476.

Key words: protein de novo design; tetrapyrrole; histidine rotamers; ligand-binding site; consensus motif.

INTRODUCTION

Heme and chlorophyll (Chl) proteins are a remarkable example of the way Nature achieves functional diversity with only a few types of cofactors (or by variations of a single type of cofactor). The proteins provide robust scaffolds for tightly binding the cofactors at well-defined geometries, as well as specific amino-acid residues that interact with the cofactors, thereby tuning their structural and electronic properties to fit the desired function. Detailed understanding of protein-cofactor interactions is critical for designing non-natural proteins with distinct activities. A useful way for gaining such understanding is by designing *de novo* simple and robust protein scaffolds with heme and Chl-binding sites. This strategy has been very successful in providing minimal analogs of a wide variety of natural heme-binding proteins.[1–5] However, there is a need for improved methods for designing heme and Chl-binding proteins that will provide higher specificity, and better control of cofactor arrangement and binding affinity.

Chls and their bacterial analogs, bacteriochlorophylls (BChls), as well as the widely used heme cofactors, are cyclic tetrapyrroles. This group includes ubiquitous cofactors of biological catalysis that are involved in a multitude of reactions. The tetrapyrrole skeleton forms an aromatic macrocycle with the pyrrole nitrogens providing four in-plane ligands. The ligated central metals differentiate the major classes: iron in the hemes and magnesium (and rarely, zinc) in (B)Chls. The latter molecules are further distinguished by

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant number: BR 1991/2-1; Grant sponsors: Human Frontiers Science Program Organization, Weizmann Institute of Science's Center for young investigators; Grant sponsors: National Science Foundation; Grant number: MCB-0920448; Grant sponsors: New York Structural Biology Center; Grant number: P41 GM-66354; Grant sponsors: NIH National Center for Research Resources; Grant number: NIH 5G12 RR03060; Grant sponsors: NIH's Minority Access to Research Careers Program; Grant number: T34 GM007639.
*Correspondence to: Dror Noy, Plant Sciences Department, Weizmann Institute of Science, Rehovot 76100, Israel. E-mail: dror.noy@weizmann.ac.il.
Received 9 July 2010; Revised 6 September 2010; Accepted 13 September 2010
Published online 7 October 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.22895

the presence of different peripheral substituents on the tetrapyrrole macrocycle as well as by its hydrogenation state. Chl and BChl are very similar: the macrocycle of BChl-a differs from that of Chl-a only by reduction of the tetrapyrrole's ring B and the replacement of a vinyl group at ring A by an acetyl group.

Axial ligation to the central magnesium atom has long been recognized as critical for the binding of (B)Chl and thus for the assembly of (B)Chl proteins.[6,7] There are, however, numerous additional possibilities of interactions with the substituents; in particular, H-bonding to the carbonyl groups seems to be widespread.[8,9] Additionally, there is the stereochemical aspect in (B)Chl ligation.[10–12] (B)Chls have several asymmetrically substituted C-atoms that render them intrinsically chiral. The two faces of the chlorin macrocycle are diastereotopic, and an additional chiral center is therefore generated by ligation of the central magnesium from either the top (α-type) or the bottom face (β-type) of the macrocycle [Scheme 1(C)]; these two ligation states should differ in their binding energies.[13,14] In natural (B)Chl proteins, the two types of ligation are unevenly distributed, and α-type ligation of (B)Chls is predominant.[9,11] β-Ligated (B)Chl has been proposed to play an important role in (B)Chl protein assembly.[9] However, the factors determining which of the two diastereotopically distinct conformations is preferred are, as yet, not known.

Designing *de novo* a protein that will fold and self-assemble with large organic cofactors such as heme or (B)Chl is a difficult challenge: the computational protein design requires solving simultaneously two difficult problems, namely, protein-folding and molecular recognition.[15,16] The latter is particularly challenging because of the numerous possibilities for binding organic cofactors that, unlike amino acid (AA) side chains, are unconstrained by any covalent link to the polypeptide backbone. To simplify the problem and improve the reliability of protein design, a variety of empirical methods have been proposed.[15–17] These rely on identifying sequence and structure motifs that are characteristic of natural cofactor-binding sites. Recently, Negron *et al.*[18] provided a set of rules for designing heme-binding sites in helical bundles. These were based on analyzing helical fragments around heme-binding histidines selected from a nonredundant database of heme-binding protein structures. The side-chain rotamers of heme-binding histidines were found to be restricted to just four of the eight global helical histidine rotamers, as identified by Lovel *et al.*[19] Moreover, distinct sequence and structural motifs were revealed for the helical fragments in each of the four histidine rotamer classes, thereby providing consensus sequences for heme-binding sites for each rotamer.[18]

Progress in heme-binding protein design is currently unmatched by similar progress in designing (B)Chl-binding proteins. This is despite the many similarities between hemes and (B)Chls in biosynthesis, chemical structure and properties, and biological function.[20] A major difficulty in binding native (B)Chls to water-soluble proteins is the marginal water solubility of the pigments. This can be improved by hydrolysis of the long-chain esterifying alcohol, resulting in (bacterio)-chlorophyllides [(B)Chlides] that are much more water-soluble and could be reconstituted into globular heme-proteins such as myoglobin and hemoglobin[21–23] as well as into *de novo* designed heme-binding proteins.[4,5,24] Another problem is the rather limited number of structurally distinct natural (B)Chl-binding proteins, compared to the wealth of heme-binding protein structures. Yet, unlike heme-proteins, the pigment content of most (B)Chl-proteins is very high; some of which, like photosystem I (PSI), maintaining dozens of different Chl-binding sites. Several strategies have been previously used for the prediction of potential (B)Chl binding and assembly motifs, including analysis of high-resolution structures, in particular, PSI, and computational analysis of the amino-acid distribution in putative (B)Chl-binding pockets of "nonhomologous" (B)Chl-proteins retrieved from protein databases.[9,25,26]

Here, we combine the previous strategies of analyzing natural (B)Chl-binding and assembly motifs, with the methodology applied by Negron *et al.*[18] for analyzing heme-binding proteins, namely, constructing a database of nonredundant helical segments with (B)Chl-binding histidines, subdividing according to histidine rotamers, and searching for distinct sequence and structural motifs. Because the structures of most of the naturally existing (B)Chl-binding proteins are currently available at medium to high resolution, it was possible to apply a more stringent set of rules in order to exclude not only redundant structures but also nonhelical (B)Chl-binding motifs while maintaining a database of considerable size. Although there are only 20 or so nonunique coordinate files of (B)Chl-proteins in the PDB, compared to more than 2000 structures of heme-proteins, we identified altogether 34 unrelated (B)Chl-bound helical histidines from six different (B)Chl-protein structures. This database is smaller but comparable to that of Negron *et al.*,[18] which contained 61 unrelated heme-bound helical histidines. Thus, we set out to identify distinct (B)Chl-binding pocket topologies as well as (B)Chl-protein and -pigment interaction patterns and thereof to derive a set of rules for (B)Chl-binding. The analysis provides useful distinction between (B)Chl and heme-binding sites, which lead to new guidelines for designing specific (B)Chl-binding helical proteins.

METHODS

Construction of the database

Database construction started by selecting the representative (B)Chl-binding protein structures from the

Table I
Nonredundant (B)Chl-Binding Chains and the Helical Histidines Database

PDB ID	Chain	Total (B)Chls (in database[a])	(B)Chl-binding helical histidine residues
1JB0	A	46 (21)	52, 76, 79, 93, 199, 200, 215, 299, 300, 301, 373, 396, 443, 461, 494, 539, 540, 547, 680, 708, 734
1JB0	F	1 (0)	
1JB0	J	2 (0)	
1JB0	K	1 (0)	
1JB0	L	3 (1)	54
2AXT	B	16 (3)	23, 100, 466
2AXT	D	2 (2)	117, 197
2BHW	A	14 (2)	68, 212
1LGH	A	2 (1)	34
1NKZ	A	2 (1)	31
1NKZ	B	1 (1)	30
2J8C	M	2 (2)	182, 202

[a]Only (B)Chls bound to helical histidines were added to the database.

PDB. The selection criteria of 70% identity cutoff, resolution of 3.0 Å or better, and the presence of at least one (B)Chl-binding helical histidine have led to the following list of PDB files (number and type of (B)Chls in each structure are in parenthesis): 1JB0 (96 Chl a), 2AXT (70 Chl a), 2BHW (24 Chl a, 18 Chl b), 1EYS (4, BChl a) 1LGH (12 BChl a), 1NKZ (9 BChl a), 2I5N (4 BChl b), and 2J8C (4 BChl a). Subsequent filtering of individual chains was done using PISCES[27] (http://dunbrack.fccc.edu/PISCES.php) with default values (identity $\leq$ 25%, resolution $\leq$ 3.0 Å, *R*-factor $\leq$ 0.3, sequence length: 40–10,000, cull by chain) and resulted in a subset of 12 nonredundant Chl-binding chains as shown in Table 1. These were scanned for (B)Chl-binding histidines within helical segments. Helicity was determined by considering the Φ/Ψ angles of the histidine ligand and additional four residues that proceed and precede it. If at least six of these nine residues, including the histidine itself, are locally helical ($-82 < \Phi < -42$ and $-61 < \Psi < -21$), then the segment was considered helical. Thus, the final helical dataset comprised 34 histidine-bound (B)Chls out of a total of 92 in the whole nonredundant subset.

Definition of structural parameters

The torsion angles, χ_1 and χ_2, and rotamer classification were determined according to Lovell *et al.*[19] Other geometric parameters, defined in Scheme 1, required alignment of the different (B)Chl-binding helical protein segments. For this purpose, the cartesian PDB coordinates of each segment and its associated (B)Chls and proteins were transformed into an internal cylindrical coordinate system (r, θ, z) whereby the z axis is aligned with the helix axis such that the helical axis is crossing $r = 0$, and Cα of the binding histidine is at $z = 0$ and $\theta = 0$. The helix axis was determined by using Kahn's method[28–30] considering the ligating histidine, five preceding, and five proceeding residues.

To mark the dimension of a (B)Chl molecule, the (B)Chl plane was defined as the one that best fits (minimum perpendicular distances) the macrocycle atoms (Scheme 1, C1–C20 and N21–N24). The (B)Chl atoms were projected once on this plane and then further onto the line formed by crossing the (B)Chl plane with the plane defined by $z = 0$ in the internal coordinate system; the projected points p_1 and p_2 that span the longest distance along this line were chosen to mark the extent of each (B)Chl molecule. For aligning (B)Chls, an internal Cartesian coordinate system was defined, whereby the X axis is the vector connecting the projections of N24 and N22 on the (B)Chl plane; the origin is the point along the X axis that intersects the vector connecting the projections of N23 and N21, and the Z axis is directed toward the viewer.

Computations of interaction energies

Calculations were carried out on a straight, 3.6 residues per turn, 30-residue polyalanine α-helix with idealized bond angles and lengths. At position 16, the center of the α-helix, a histidine bound to (B)Chl group, was incorporated using exemplar coordinates from the nonredundant database. These included HIS-A540/CL1-1136 (m83-rotamer, Nε ligand) and HIS-A396/CL1-1126 (t78-rotamer, Nδ ligand) from PDB file 1JB0 and HIS-B466/CLA-1136 (t60-rotamer, Nε ligand) from PDB file 2AXT.

With the exception of glycine and proline, the energy of interaction between each type of AA and the His-(B)Chl complex was computed using a 12-6 Lennard Jones function with atomic parameters from the AMBER 94 force field[31] as implemented in protCAD.[32] This only takes into account packing interactions and does not consider electrostatics or hydrogen bonding contributions to stability.

The reported energy for AA type i at position j relative to the (B)Chl-binding histidine was computed as

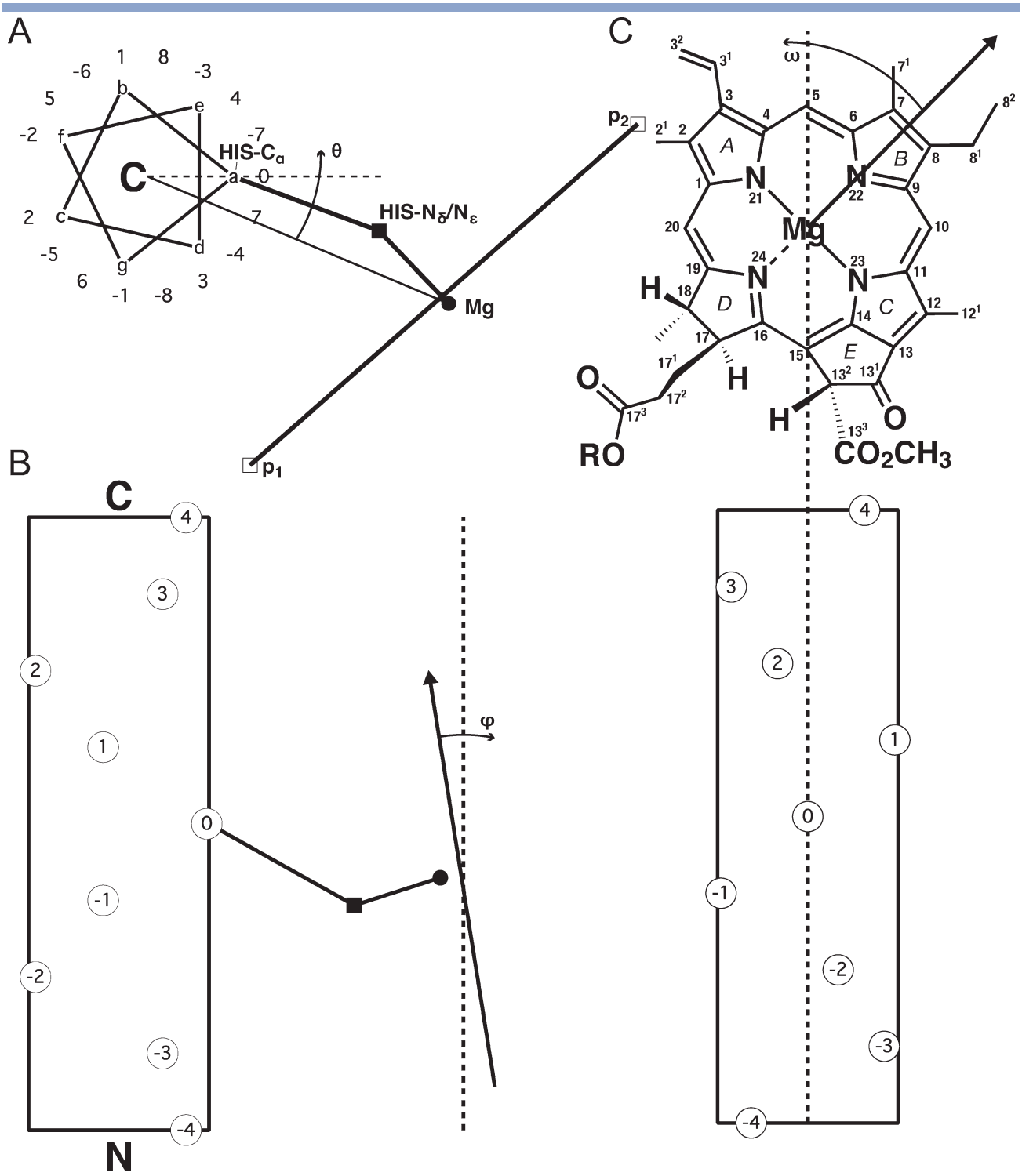
$$E(i,j) = \min_k E(i,j,k) - E(template),$$

where k was the set of most frequent discrete rotamers from the BBDEP rotamer library.[33] E(template) was the energy where i = alanine, which was constant across all cases. Energies were calculated for ($-8 \leq j \leq -1$) and ($1 \leq j \leq 7$). To compute position specific energies for alanine itself, polyglycine was used in place of a polyalanine template.

RESULTS AND DISCUSSION

Histidine ligand rotamers and (B)Chl topologies

The torsion angles distribution of the (B)Chl-binding histidines in the nonredundant helical dataset are depicted in Figure 1(A) and summarized in Table 2. The



Scheme 1

Structural parameters of (B)Chl-bound helical histidines. **A:** The Cartesian PDB coordinates of (B)Chls and proteins were transformed into an internal cylindrical coordinate system (r, θ, z), whereby the z axis is aligned with the helix axis such that the helical axis is crossing $r = 0$, and Cα of the binding histidine is at $z = 0$ and θ = 0. The helix is represented as an ideal heptad repeat with seven distinct position labeled $a - e$ and the binding histidine at position a. The angular positions of residue −8 to +8 relative to the binding histidine are also noted based on 18 points helical wheel representation. The (B)Chl is represented by the central Mg atom (solid circle) and a projection of a line drawn on the (B)Chl plane between points p_1 and p_2 (see Methods section for a detailed definition). The histidine's nitrogen ligand is indicated by the solid square. **B:** The tilt angle φ is the angle between the helix axis and the (B)Chl plane. **C:** The rotational angle ω is the counter-clockwise angle between a vector connecting the projection of (B)Chl atoms N_{24} and N_{22} and the helical axis. The Chl atoms considered in this work are shown labeled according IUPAC conventions. BChl atoms are identical except for O_3^2 replacing $C3^2$ and the reduced ring B. The phytyl chains (R) are not considered. In α- and β-ligated (B)Chls, the histidine ligand and $C17^1$ of the propionic acid side chain are either on opposite or on the same sides of the (B)Chl plane, respectively.

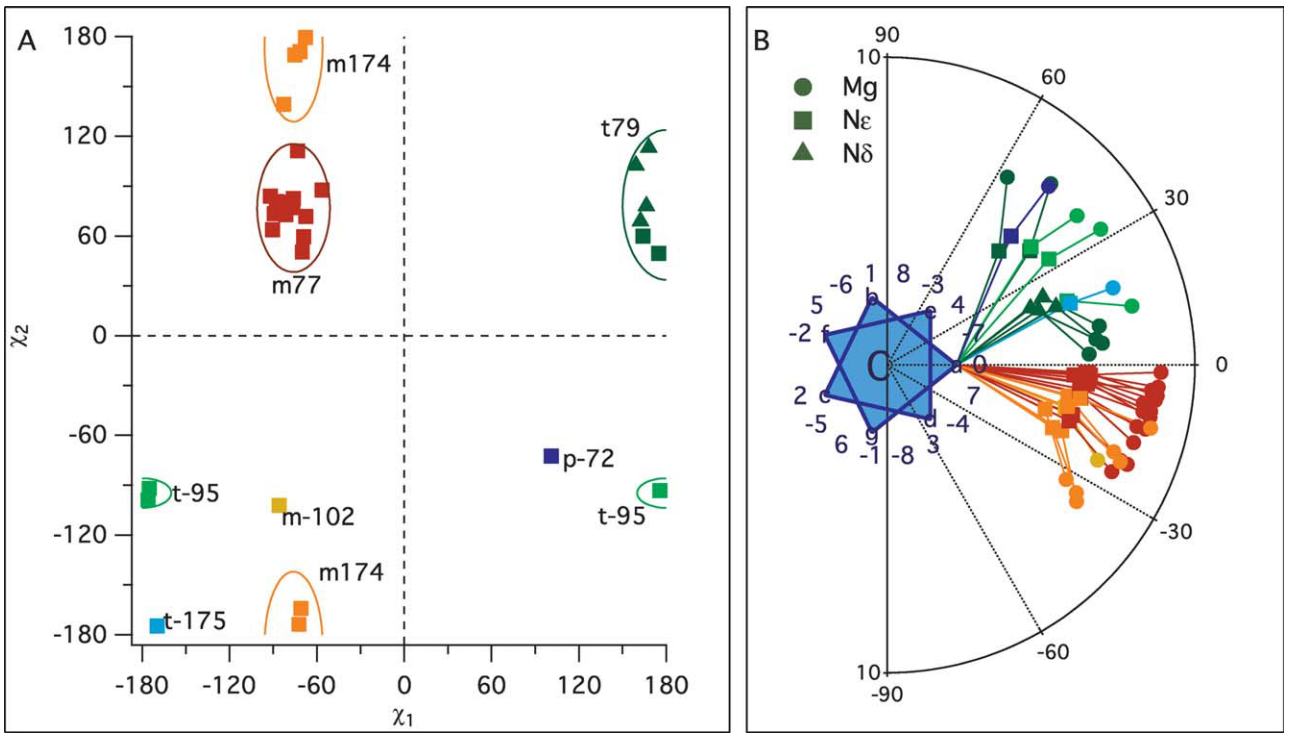


Figure 1

A: Distribution of the side-chain torsion angles, χ_1 and χ_2, in the nonredundant dataset of (B)Chl-bound helical histidines. **B:** Angular and radial positions of histidine ligands and the Mg atoms of (B)Chls color coded according to their respective rotamers. The coordinating Nε and Nδ nitrogens of each histidine and the Mg atoms of the (B)Chls are marked by squares, triangles, and circles, respectively.

distribution clusters into four rotamers, m77, m174, t79, and t95, equivalent to the m80, m166, t73, and t86 rotamers, respectively, observed by Negron *et al.*[18] in the dataset of helical heme-binding histidines. Additionally, three histidines are each a single example of the m102, t175, and p72 rotamers, equivalent to the m70, t166, and p80 histidine rotamers defined by Lovell *et al.*[19] Strikingly, the distributions of the heme- and (B)Chl-binding histidines are very different. The m-rotamers are predominant in (B)Chl-binding histidines and make up 67.5% of the whole dataset, with the m77 being the predominant rotamer making up 47% of the dataset. Conversely, t-rotamers make up 53% of the heme bound histidines dataset, with the t73 rotamers making up 41% of the dataset.[18] Moreover, of 22 bis-His ligated hemes, 16 (73%) were found to be ligated by pairs of t73 histidines.

Table II
Rotamer Distributions of (B)Chl-Binding Helical Histidines Compared to Heme- and Nonbinding Histidines

(B)Chl-binding helical histidines[a]				Heme-binding helical histidines[b]				All helical histidines[c]	
Rotamer	Count (%)	χ_1 (SD)	χ_2 (SD)	Rotamer	Count (%)	χ_1 (SD)	χ_2 (SD)	Rotamer	Occurence (%)
m77	16 (47)	−79 (10)	77 (13)	m80	12 (20)	−79 (6)	80 (14)	m80	14
m174	6 (17.5)	−73 (5)	174 (20)	m166	16 (27)	−73 (7)	166 (18)	m170	9
m-102	1 (3)	−86	−102	—	—	—	—	m-70	26
Total m		23 (67.5%)		Total m		28 (47%)		Total m	49
t79	6 (17.5)	166 (5)	79 (25)	t73	24 (41)	−173 (13)	73 (15)	t60	24
t-95	3 (9)	−178 (5)	−95 (4)	t-86	7 (12)	−164 (9)	−86 (11)	t-80	17
t-175	1 (3)	−170	−175	—	—	—	—	t-160	5
Total t		10 (29.5%)		Total t		31 (53%)		Total t	46
p-72	1 (3)	101	−72	—	—	—	—	p-80	0
Total p		1 (3%)		Total p		0 (0%)		Total p	0

[a]This work.
[b]Negron *et al.*[18]
[c]Lovell *et al.*[19]

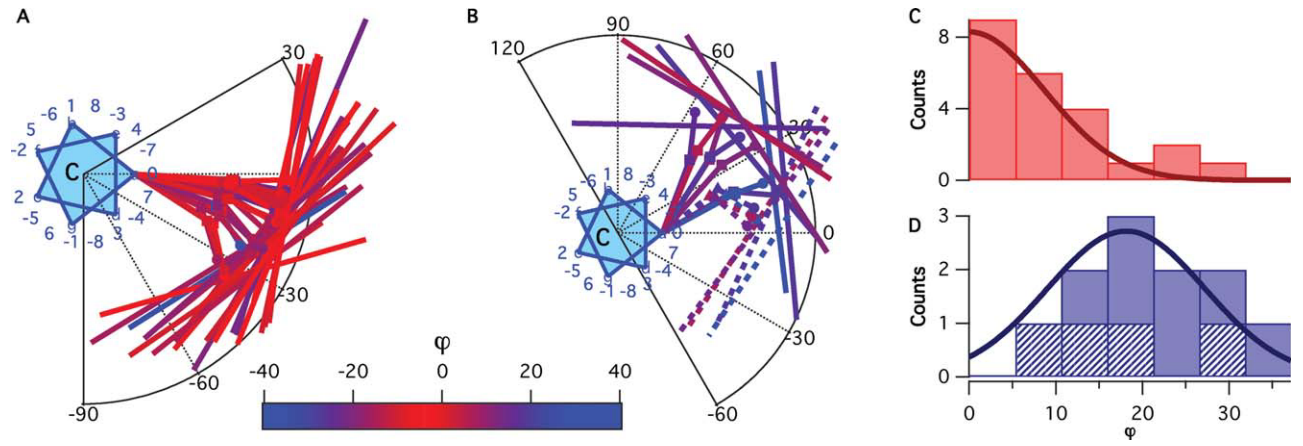


Figure 2

Macrocycle topology of m- and t-type rotamer-ligated (B)Chls. Angular and radial positions of m- (**A**) and t- (**B**) rotamer histidines and their bound (B)Chls colored according to the tilt angle, φ, of the (B)Chl plane. Solid lines are used for Nϵ-bound and dotted lines for Nδ-bound (B)Chls. Distributions of φ absolute values in m- (**C**, red bars) and t- (**D**, solid blue bars represent Nϵ-bound, striped bars Nδ-bound (B)Chls) rotamers are centered about $|\varphi| = 0^\circ$ and $|\varphi| = 19^\circ$, respectively.

The differences in histidine rotameric conformation significantly affect the position of the hemes and (B)Chls with respect to the binding helical axis. As shown in Figure 1(B), the Nε nitrogens of m-rotamer histidines are located at azimuthal angles between 0° and 20°, and the corresponding central Mg atoms of bound (B)Chls are between 0° and −36°. Respectively, the Nε nitrogens of t-rotamer histidines are between 20° and 46°, and the corresponding central Mg atoms of bound (B)Chls are between 14° and 56°. Notably, four t-rotamer histidines bind a (B)Chl at their Nδ nitrogen, which brings the central Mg atoms closer to the binding helix at azimuthal angles between 0° and 10°. In contrast to the situation with heme-binding histidines, metal coordination to the Nε nitrogens is not exclusive in (B)Chl-binding histidines, although it is considerably preferred over the coordination to the Nδ nitrogens. This reflects the general trend of His-Fe and His-Mg coordination as demonstrated by Chakrabarty.[34]

The different positions of the central Mg atoms (and their histidine ligands) obviously correlate with distinct topologies, orientations, and conformations of the (B)Chl macrocycles in the helical-binding pocket. Particularly, the helix-(B)Chl interfaces are clearly distinct, that is, residues at distinct positions in the helix register interact with either the m- or t-rotamer bound (B)Chls [Fig. 2(A,B)]. The m-rotamer bound (B)Chls are close to the residues at positions −8, ±7, −4, −1, and +3 (relative to the ligand histidine at position 0), whereas the t-rotamer bound (B)Chls are close to the residues at +8, ±7, −6, ±4, ±3, and +1. In addition, the (B)Chls bound to m- and t-rotamer histidines have different tilt angle distributions [Fig. 2(C,D)]. Specifically, most of the m-rotamer bound (B)Chls have small tilt angles (more than two-third have tilt angles smaller than 10°), hence their macrocycles are almost parallel to the helix axis [Fig. 2(C)]. In contrast, the t-rotamer bound (B)Chls are generally more tilted (two-third of the (B)Chl have tilt angles larger than 10°) with respect to the helical axis [Fig. 2(D)].

The rotational angles of (B)Chls with respect to the binding helical axis are shown in Figure 3. Conspicuously, the (B)Chls' rotation angle distribution is not continuous but clustered in three distinct orientations centered around −153°, −19°, and 95°. However, in contrast to the tilt angle distribution that clearly distinguishes between (B)Chls bound to either m- or t-rotamer histidines [Fig. 2(C,D)], there is no obvious dependence of the rotational angle on the type of histidine rotamers. This is similar to the findings of Fufezan *et al.*[35] indicating that rotation angles of histidine ligand to noncovalently bound (b-type) hemes are evenly distributed. Instead, the rotational angles appear to be dependent on the (B)Chl ligation mode. Notably, rotational angles centered around 95° are exclusively found among the α-ligated (B)Chls (see Fig. 3). This observation may be explained by considering the steric hindrance of the bulky phytyl chains that are found on the binding helix side of β-ligated (B)Chls and across it in α-ligated (B)Chls. The phytyl chains are on the right side of the binding helix (viewed from the (B)Chl direction) in the −153° and −19° clusters and on the left side in the 95° cluster. Because the left side is usually closer to the binding helix (see Fig. 2), clashes with the phytyl chains are more likely in the β-ligation mode making this conformation unfavorable.

Taken together, the analysis of macrocycle orientations in the nonredundant dataset of (B)Chl-bound helical

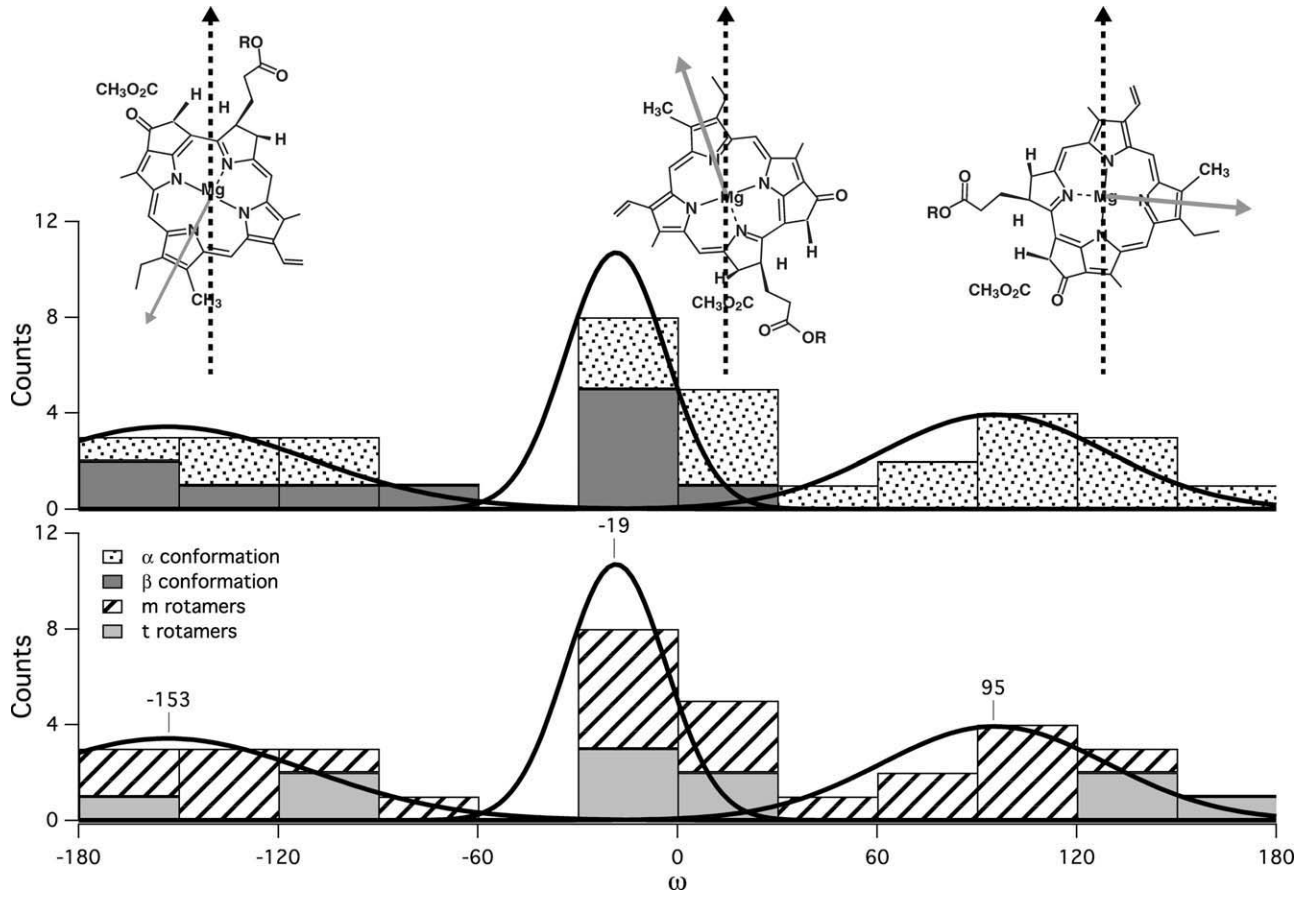


Figure 3

Rotational angle (ω) distribution of (B)Chls is distinctly discontinuous and depend on their (B)Chl α- or β-ligation mode (top histogram, dotted, and dark gray bars, respectively) but not on the m- or t-rotamers of binding histidines (bottom histogram, striped, and light gray bars, respectively). Notably, the 95° rotation of (B)Chl is found only in α conformers.

histidines indicates distinct (B)Chl-helix topologies strictly dependent on histidine rotamer (primarily m or t) as well as on (B)Chl ligation mode (α or β). To explore whether particular interactions between (B)Chl and specific AA residues accompany the specificity of the binding modes, we have rigorously analyzed the contacts between the (B)Chl macrocycles and their binding helices.

(B)Chl interactions with the binding helix

The patterns of interactions observed between the (B)Chls and their binding helices (see Fig. 4) are remarkably consistent with the proposed interaction patterns derived from the histidine rotamer-based topologies of the (B)Chl macrocycles [Fig. 2(A,B)]. Thus, (B)Chls bound to m-rotamer histidines [Fig. 4(A)] interact primarily with residues at positions −4, −1, and +3, less frequently with positions ±8, −5, +4, and ±7, but never with positions ±6, −3, ±2, 1, and 5. (B)Chls bound to Nε of t-rotamer histidines [Fig. 4(C)] interact primarily with residues at position +4, less frequently with positions −6, −5, ±3, 1, 7, and 8, but never with positions −8, −4, ±2, −1, 5, and 6. The interaction pattern of (B)Chls bound to Nδ of t-rotamer histidines is a combination of t- and m-rotamer patterns, as may be expected from their macrocycle orientation.

The interaction patterns identified here were obtained from a limited set of protein structures. Extending the database by adding homologous sequences from different species is, however, impractical as the AA sequences of the (B)Chl-binding proteins are highly conserved among the various species. To validate our findings, we modeled the energetic states of specific AA at the critical contact points in the (B)Chl-binding pockets with the computational protein design package, protCAD. This software, like other protein design programs, uses molecular mechanics potentials, knowledge-based potentials, or a weighted combination of both in optimizing the AA sequence for a target fold. Using a straightforward van der Waals potential that scored for atomic packing, the energies of eighteen AA were determined through one

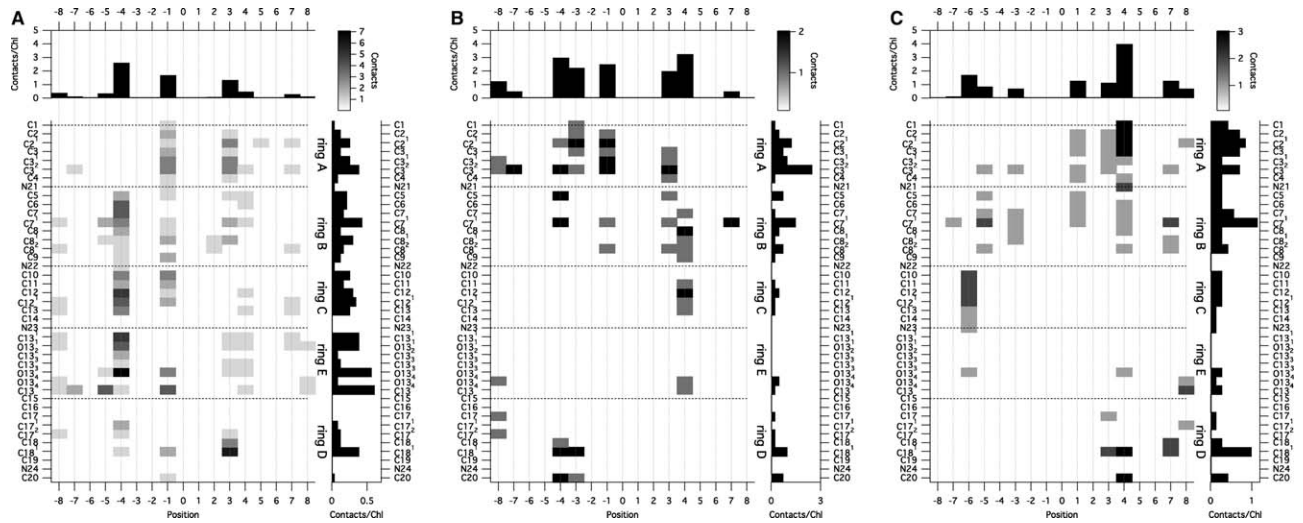


Figure 4

Contact maps of (B)Chl and AA residues along the binding helix [positions ±8 relative to ligating histidine: (**A**) m-rotamers, (**B**) t-rotamers with Nδ as ligand, and (**C**) t-rotamers with Nε as ligand]. The color coding indicates contact frequency.

full helical turn on either side of the coordinating histidine (see Fig. 5). Glycine and proline were not included in the current calculation due to their strong effects on the helix stability. Position energies were calculated in the presence of a histidine-bound Chl at position 0 and fixed in either the m or t-rotamer with all other positions set to alanine. Reference energies were calculated using a polyalanine helix devoid of histidine and Chl. The only exception was the energy profile for alanine itself, which was computed in a polyglycine helix background. As shown in Figure 5, the computed position energies that are most affected by the presence of Chl match those positions that were found to interact more frequently with the (B)Chl macrocycle (see Fig. 4). These results are also in agreement with computational studies by Negron *et al.*[18] of a model octamethyl porphyrin bound to a single histidine flanked by alanines in an ideal α-helix. These studies found nearly identical patterns of interactions between the protein residues and either m- or t-rotamer-ligated porphyrins.

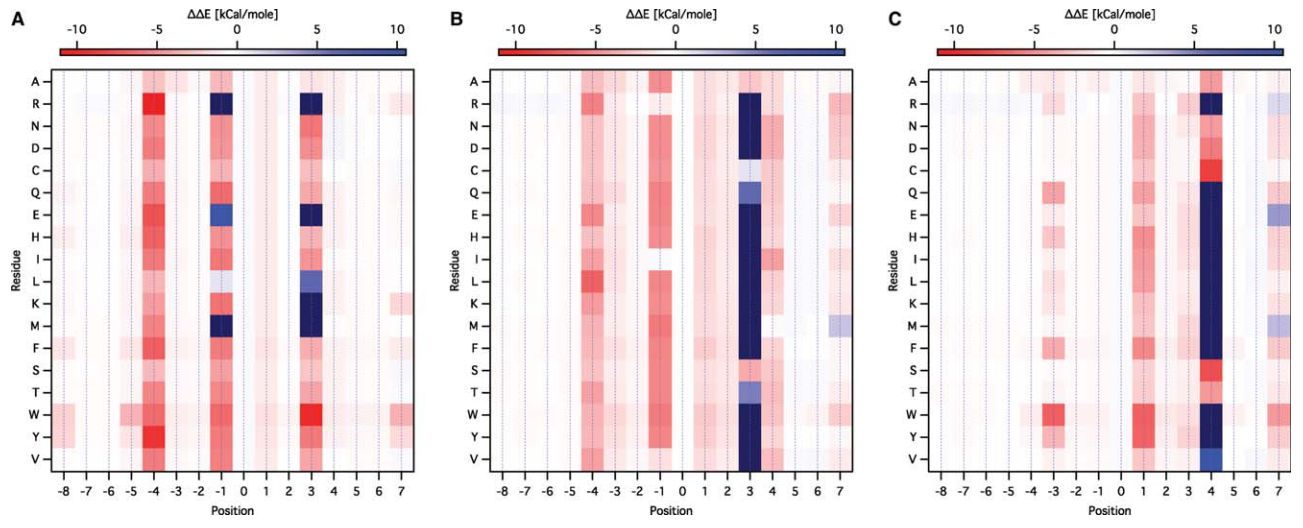


Figure 5

Relative energy ($\Delta\Delta E$) maps versus position and amino acid type: (**A**) m-rotamers, (**B**) t-rotamers with Nδ as ligand, and (**C**) t-rotamers with Nε as ligand. $\Delta\Delta E$ values larger than 10 kcal/mol are colored dark blue.

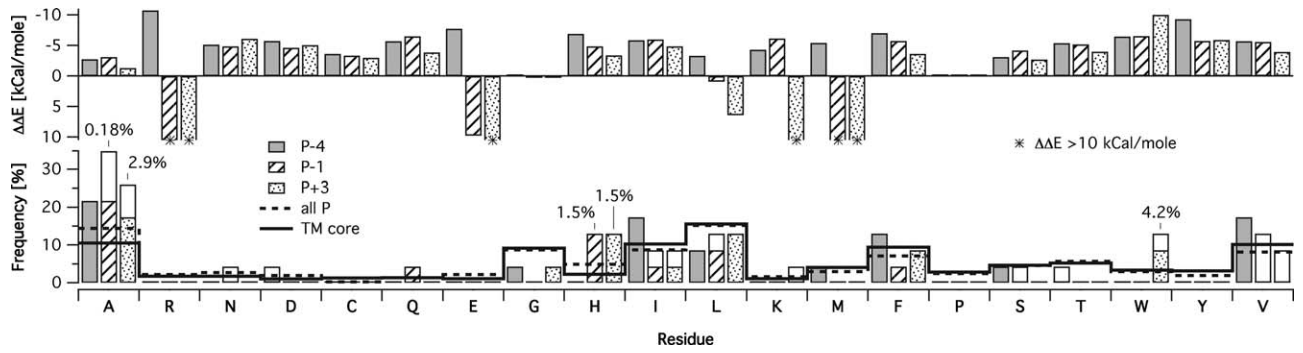


Figure 6

Amino acids frequencies (bottom) and energies calculated by protCAD (top) at positions $P-4$ (light gray), $P-1$ (striped), and $P+3$ (dotted) where $P0$ is the position of m-rotamer Chl-binding histidine. Empty bars represent occurrence frequency of residues at a given position, and full bars represent interaction frequency, that is, only residues that interact with the bound (B)Chl. Black and red lines indicate relative amino acid frequencies in TM core of a random data set and those of (B)Chl-binding proteins of the data set used in this study, respectively. Probability values are marked for amino acids that scored below the 5% significance level in a two-sided binomial test versus the TM core frequencies.

As shown, there are distinct BChl–protein interaction motifs correlated with the rotamers of binding histidines. In m-rotamer bound (B)Chls, the atoms of the pyrrole rings B, C, and E are preferentially in close contact with the residues at position −4, while the atoms of rings A and D are preferentially in contact with the residues at position +3, and the residues at position −1 can be in contact with any atom of the entire macrocyle [Fig. 4(A)]. Most noticeably, the $C13^3$ oxo and the $C18^1$ carbons of the m-rotamer bound (B)Chls have the highest contact frequencies and interact preferentially with the residues at −4 and +3, respectively. This interaction pattern is compatible with the topology of the (B)Chls oriented either around 95° or −153° with respect to the binding helix but not with those oriented around −19° (see Fig. 3). Likewise, there are distinct patterns emerging for the t-rotamer BChls [Fig. 4(B,C)]. In contrast to the m-rotamer (B)Chls, the (B)Chl-helix contacts of the t-rotamer (B)Chl are preferentially with the pyrrole rings A, B, and D while rings C and E have clearly less contacts with the (B)Chl. However, due to the limited overall number of t-rotamers bound to (B)Chls, which again are further split into Nε- and Nδ-bound species, additional samples are necessary to validate this finding. Nevertheless, as obvious from the (B)Chl-protein contact maps, distinct rotamer-dependent (B)Chl-helix interaction patterns exist, which suggests that the (B)Chl molecules assume distinct topologies in their binding pockets. This is further supported by the results obtained from the analyses of histidine-bound (B)Chls conformations showing that the (B)Chl macrocycles assume distinct tilt and rotational angles, with respect to the binding helix (Figs. 2 and 3). The topologies of the (B)Chl are likely stabilized by packing against the helix interface and specific protein-(B)Chl interactions. To identify such interaction motifs and to test whether there are particular residues at structurally critical positions of (B)Chl-protein contact, the amino-acid distribution at these positions has been examined.

Consensus sequences for (B)Chl-binding motifs

The AA distributions at positions −4, −1, and +3 of the m-rotamer subset were compared to their distributions in the entire subset as well as in the transmembranal (TM) core of a "random" set of helical TM proteins. The latter is based on a recent statistical analysis of a nonredundant set of TM helical structures[36] (see Fig. 6). The AA residues appear very similar, and this observation was validated statistically by a χ^2 test comparing the two distributions. The test yielded $\chi^2 = 4.5$, corresponding to a probability (P) value of 99.97% that the two distributions are identical. The frequencies of occurrence at positions −4, −1, and +3 seem to follow the trend of core TM distribution with a few clear deviations. However, these should be treated with some caution considering that the total number of counts in each position is only 23. Thus, we used a two-sided binomial test for comparing the occurrence of each AA at each position to its expected frequency according to the apparently random TM core distribution. Three residues, namely alanine and histidine at positions −1 and +3 and tryptophan at position +3, were found to be significantly overrepresented according to the binomial test ($P < 5\%$).

The occurrence of histidine can be rationalized by its specific role as a (B)Chl ligand and the fact that many of the helices in the dataset bind more than one (B)Chl. Generally, the occurrence of histidine (as well as other charged and polar residues) in TM proteins is restricted to specific functional motifs. In this context, the absence of histidine from position −4 is noteworthy, although

Position	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
t 60 (protCAD)	X	X	X	X	X	QFYW	X	X	**H**	HIFYW	X	X	ANDST	X	X	X
heme t73	I	I	X	X	X	F	X	X	**H**	A	X	G	I	I	I	F
m 80 (protCAD)	X	X	X	X	REFYWH	X	X	QIKW	**H**	X	X	NDYW	X	X	X	W
all m rotamers	X	X	X	X	AH	X	X	AH	**H**	X	X	W	X	X	X	X

Scheme 2

Comparison between consensus sequences of (B)Chl-binding helices that were computationally derived by the protCAD software and those derived from nonredundant databases of (B)Chl- and heme-binding structural motifs.

our experimental data showed that histidine may, in principle, be placed at position $i - 4$.[37] Apparently, it hints at a functional/structurally disfavored configuration of two closely spaced (B)Chls on the same helix interface. Alanine is significantly overrepresented in positions −1 and +3, regardless of interactions with (B)Chls. This implies contributions of alanines at these positions to the overall stability of the protein fold. Interestingly, although impossible to verify statistically because of sample size, the rare occurrence of glycine in all critical positions and its complete exclusion from position −1 is noteworthy and consistent with our previous experimental data.[37,38] Our experiments revealed that placing glycine at position −1 in natural and model (B)Chl protein complexes significantly destabilized the complex, while other residues showed little effect. One explanation for this glycine-specific effect at the (B)Chl-protein interface may be that the planar structure of the phorphyrin core is devoid of groups that could fill the void created by the absence of any side chain.

Given the limitations of the small database, we used energies computed by the protCAD software [Fig. 5(A)] to gain more insight into position specific AA preferences. Clear correspondence was found between computed favorable energies and Chl interaction frequency with alanine at positions −4, −1, and +3, isoleucine, valine, and phenylalanine at position −4, and tryptophan at position +3. For others such as tyrosine at positions −1 and +3, as well as small or charged AA such as serine or aspartate, the low counts in the database made it difficult to accurately assess the correspondence. Similarly, cases of poor correlation between observed and computed preferences were presumably due either to the limitations encountered by the sparse data set or by the assumptions that have been made for the calculations such as discrete AA rotamers, an idealized helical backbone, and ignoring solvation and electrostatics contributions in the energy.[39] Altogether, the sequence analysis of the helical (B)Chl-binding pockets with m-rotamer histidine ligands and protCAD calculated interaction energies reveal characteristic position-dependent patterns of AA. The consensus sequences, H/AXXH/AHXXW, for a (B)Chl-binding α-helix with an m-rotamer histidine ligand, are derived from the analysis of the nonredundant databases of (B)Chl-binding structural motifs. The computational and database-derived consensus sequences are generally in agreement (Scheme 2). Both are significantly different from the consensus sequences identified by Negron *et al.*[18] for heme-binding helical proteins. Thus, implanting the (B)Chl-binding sequences as identified within a helical-bundle protein template may promote preferential binding of (B)Chl derivatives compared to hemes.

Chl interactions with the protein and pigment environment

Unlike hemes that are usually ligated from both sides of their macrocycle plane, (B)Chls are ligated only from one side. Additionally, in the crowded arrangements of pigments required for light-harvesting functionality, inter-(B)Chls interactions as well as interactions with other photosynthetic pigments become a considerable structural factor. The inherent chirality of (B)Chls adds an extra level of complexity to the description of the (B)Chl's protein/pigment environment, because the two diastereoscopically distinct sides of the macrocycle plane imply a clear division between ligated and nonligated sides as well as between α- and β-ligated (B)Chls. In our

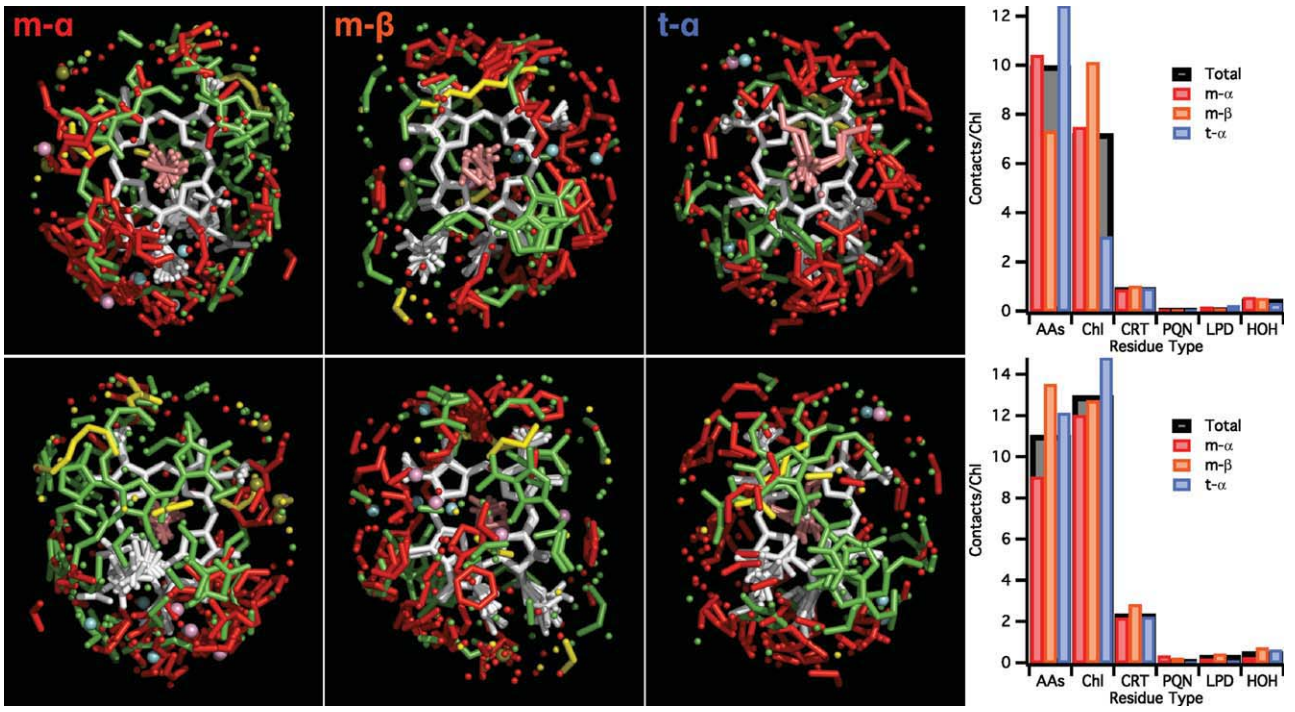


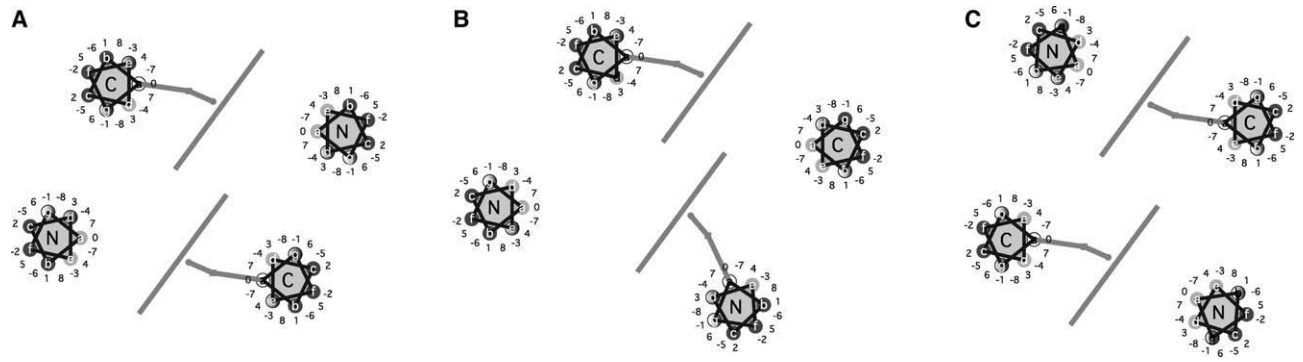
Figure 7

Interactions of helical histidine-bound (B)Chls with pigment, protein, and solvent environment. Bound (B)Chls are aligned as described in Methods section. Views and statistics are shown for interacting residues on the same side of the binding helix (top panels) and across it (bottom panels). (B)Chls α- and β-ligated to m-rotamer histidines (m-α and m-β, respectively), and (B)Chls α-ligated to t-rotamer histidines (t-α) are considered. Protein-, (B)Chl, and carotenoid residues within 4 Å from the bound (B)Chl are shown in red, green, and yellow stick and small spheres representation, respectively, except ligating histidines shown in pink. Water, lipid, and quinone atoms are shown as blue, pink, and golden spheres, respectively. The respective distributions of contacts per bound (B)Chl are shown to the right of each panel. Residue types are amino acids (AAs), neighboring (B)Chls (Chl), carotenes (CRT), lipids or detergent molecules (LPD), and water (HOH). The distribution for all (B)Chls in the database is shown as gray bars in the background.

database, 13 (B)Chls are α-ligated to m-rotamer histidines and 10 to t-rotamer histidines (m–α and t–α, respectively), 10 others are β-ligated to m-rotamer histidines (m–β), and only one is β-ligated to a t-rotamer histidine (t–β). Figure 7 provides a three-dimensional view of the complicated and variable environment of (B)Chls by aligning all and presenting the residues within four Engstroms of each (B)Chl.

By counting interactions per bound (B)Chl according to residue type, we find major and equal contributions from protein residues (excluding residues from the binding helix) and neighboring (B)Chls; carotenoids make minor contributions, whereas the contribution of other molecules such as water, lipids, and quinones is marginal. The specific distributions of interacting residues according to different rotamers, ligation modes, and the ligated and nonligated sides of the macrocycle reflect steric characteristics of each binding mode. Thus, the nonligated side of the (B)Chl is in contact with a similar number of protein and neighboring (B)Chl residues (with a slight preference for the latter), whereas the ligated side of the macrocycle has significantly more contacts with proteins than with (B)Chl residues. Furthermore, while the contact distributions at the ligand side vary significantly with ligation mode, the distributions at the nonligated side follow the same trend in each of the three ligation modes (except a slight deviation in the m–β mode). These observations are not surprising considering that the ligand side is closer to the binding protein.

Notably, the distribution of m–α (B)Chls contacts from the ligand side follows the overall trend, whereas t–α (B)Chls have significantly less contacts to neighboring (B)Chls and more contacts to protein residues compared to the overall trend. These deviations may be explained by the closer proximity of the binding helix to the (B)Chl macrocycle in t rotamers compared to m-rotamers (see Fig. 2), making it harder to position other (B)Chls in the near vicinity of the bound (B)Chl. The m–β (B)Chls have significantly more contacts with neighboring (B)Chls from the ligand side, which probably reflects the formation of β–β (B)Chl dimers. These were recently shown by Balaban *et al.*[40] to be particularly significant in the networks of



Scheme 3

Organization schemes for single-chain four-helix bundles incorporating a dimer of (B)Chls, assuming that (B)Chls are bound to the m80 rotamer of histidine. Possible configurations include (B)Chls interacting at their nonligated sides and bound between antiparallel (**A**) or parallel (**B**) helices as well as (B)Chls interacting at their ligated sides and bound between antiparallel helices (**C**). Helices are shown as ideal heptade repeats with seven distinct position labeled *a* − *e* and the binding histidine at position *a*. The angular positions of residue −8 to +8 with the binding histidine at position 0 are also noted based on 18 points helical wheel representation. The notations "C" or "N" indicating whether the carboxy- or amino-end is closer to the viewer. Positions with hydrophilic and hydrophobic residues are marked by dark and light gray background, respectively. Intermediate positions are shown with black to white-color gradient.

excitation energy transfer. β–β (B)Chl dimers are more tightly packed compared to α–β and α–α dimers. The bulky phytyl chain that faces the binding helix side keeps the bound (B)Chl macrocycle slightly away from the helix, which provides room for interactions with neighboring (B)Chl. Particularly, rings C and E are accessible to interactions from the ligand side in β-ligation mode, because the bulky C-13^2 acetyl group is on the opposite side.

Implications for designing Chl-binding helical proteins

Our survey of natural (B)Chl-protein structures and its comparison to a similar survey of natural heme-binding proteins[18] reveal that despite the significant chemical homology between heme and (B)Chls, there are distinct binding motifs for each of the two cofactor classes, in particular, concerning histidine ligand rotamer and the AA residue distribution in the respective binding pockets. These findings have important implications to the design of specific heme- or (B)Chl-binding sites in water-soluble helical bundles. The preferred conformation of (B)Chl-binding histidines is the m-rotamer, whereas heme-binding histidines prefer the t-rotamer. This preference is critical for conferring either heme- or (B)Chl-binding selectivity to helical bundles, because the rotamer type defines the azimuthal angle of the ligating nitrogen, and hence the relative orientation of the bound cofactor with respect to the helix axis (Figs. 1 and 2). Consequently, distinct binary patterns[15,16,41] of hydrophobic and hydrophilic residues are required for each rotamer to form a proper hydrophobic core that will accommodate the large macrocycles of either heme or (B)Chl.

Typically, each of the bundle's individual helices has hydrophobic residues lining one face and hydrophilic residues lining the other; bundles are thus assembled by forming a hydrophobic helix–helix interface at their core with the remaining hydrophilic residues forming the bundle's exterior. Binding heme or (B)Chl requires packing the hydrophobic tetrapyrrole macrocycle within the bundle's core. Our findings set guidelines for positioning the hydrophobic residues along the helix register according to (B)Chl macrocycle orientation and histidine rotamer. For example, having a t-rotamer histidine ligand configuration at position *a* implies that residues *a*, *b*, and *e* will form the bundle's hydrophobic core, and position *d* will be interfacial. Conversely, if the histidine at position *a* is m-rotamer, then positions *a*, *d*, and *g* will form the core, and position *e* will be interfacial.

An additional complication in designing (B)Chl proteins arises from asymmetric substituents of (B)Chl and its tendency to bind only one axial ligand. This makes coordination from one side of the macrocycle different from the other side, because the two sides are diastereotopically distinct. Thus, inherently asymmetric designs of single-chain helical bundles should be beneficial for (B)Chl-binding. Recent progress in protein computational design has made such a design feasible; a few publicly available software packages such as protCAD[32] and rosettaDesign[42] provide the necessary facilities to handle this challenging task. The recent computational design of an asymmetric four-helix bundle that binds two synthetic zinc–porphyrin derivatives is a notable example.[43] The insights gained in this work about the preferred configurations of helix-bound (B)Chls provide useful guidelines for the construction

of preliminary templates that can be optimized by computational tools.

As an example, we consider the design of a four-helix bundle binding a (B)Chl dimer (Scheme 3). Assuming the two histidines that bind the (B)Chl dimer are both m80 rotamers, different binary patterns are required for each of the four helices to form an optimal hydrophobic core. These patterns will depend on the specific configuration of the (B)Chl dimer as depicted in Scheme 3. In the next step, the binary patterns, as well as consensus sequences for specific positions along the binding helix, will be used as constraints to the computational algorithm. We intend to test such designs in the future work.

REFERENCES

1. Ghirlanda G, Osyczka A, Liu W, Antolovich M, Smith KM, Dutton PL, Wand AJ, DeGrado WF. De novo design of a D2-symmetrical protein that reproduces the diheme four-helix bundle in cytochrome bc1. J Am Chem Soc 2004;126:8141–8147.
2. Gibney BR, Isogai Y, Rabanal F, Reddy KS, Grosset AM, Moser CC, Dutton PL. Self-assembly of heme A and heme B in a designed four-helix bundle: implications for a cytochrome c oxidase maquette. Biochemistry 2000;39:11041–11049.
3. Huang SS, Koder RL, Lewis M, Wand AJ, Dutton PL. The HP-1 maquette: from an apoprotein to a structured hemoprotein designed to promote redox-coupled proton exchange. Proc Natl Acad Sci USA 2004;101:5536–5541.
4. Koder RL, Valentine KG, Cerda J, Noy D, Smith KM, Wand AJ, Dutton PL. Nativelike structure in designed four α-helix bundles driven by buried polar interactions. J Am Chem Soc 2006;128:14450–14451.
5. Razeghifard AR, Wydrzynski T. Binding of Zn-chlorin to a synthetic four-helix bundle peptide through histidine ligation. Biochemistry 2003;42:1024–1030.
6. Coleman WJ, Youvan DC. Spectroscopic analysis of genetically modified photosynthetic reaction centers. Annu Rev Biophys Biol 1990;19:333–367.
7. Olsen JD, Sturgis JN, Westerhuis WHJ, Fowler GJS, Hunter CN, Robert B. Site-directed modification of the ligands to the bacteriochlorophylls of the light-harvesting LH1 and LH2 complexes of *Rhodobacter sphaeroides*. Biochemistry 1997;36:12625–12632.
8. Kwa LG, Garcia-Martin A, Vegh AP, Strohmann B, Robert B, Braun P. Hydrogen bonding in a model bacteriochlorophyll-binding site drives assembly of light harvesting complex. J Biol Chem 2004;279:15067–15075.
9. Garcia-Martin A, Kwa LG, Strohmann B, Robert B, Holzwart A, Braun P. Structural role of bacteriochlorophyll ligated in the energetically unfavourable β-position in light harvesting complex. J Bio Chem 2006;281:10626–1034.
10. Oba T, Tamiaki H. Effects of peripheral substituents on diastereoselectivity of the fifth ligand-binding to chlorophylls, and nomenclature of the asymmetric axial coordination sites. Bioorg Med Chem 2005;13:5733–5739.
11. Balaban TS. Are syn-ligated (bacterio)chlorophyll dimers energetic traps in light-harvesting systems? FEBS Lett 2003;545:97–102.
12. Balaban TS, Fromme P, Holzwarth AR, Krauss N, Prokhorenko VI. Relevance of the diastereotopic ligation of magnesium atoms of chlorophylls in Photosystem I. Biochim Biophys Acta Bioenerg 2002;1556:197–207.
13. Oba T, Tamiaki H. Which side of the pi-macrocycle plane of (bacterio)chlorophylls is favored for binding of the fifth ligand? Photosynth Res 2002;74:1–10.
14. Kania A, Fiedor L. Steric control of bacteriochlorophyll ligation. J Am Chem Soc 2006;128:454–458.
15. Koder RL, Dutton PL. Intelligent design: the de novo engineering of proteins with specified functions. Dalton Trans 2006:3045–3051.
16. Nanda V, Koder RL. Designing artificial enzymes by intuition and computation. Nat Chem 2010;2:15–24.
17. Moffet DA, Hecht MH. De novo proteins from combinatorial libraries. Chem Rev 2001;101:3191–3203.
18. Negron C, Fufezan C, Koder RL. Geometric constraints for porphyrin binding in helical protein binding sites. Proteins: Struct Funct Bioinf 2009;74:400–416.
19. Lovell SC, Word JM, Richardson JS, Richardson DC. The penultimate rotamer library. Proteins-Struct Funct Genet 2000;40:389–408.
20. Haehnel W, Noy D, Scheer H. De novo designed bacteriochlorophyll-binding helix-bundle proteins. In: Hunter CN, Daldal F, M.C. T, Beatty JT, editors. The purple phototrophic bacteria. Dorderecht: Springer; 2008. pp895–912.
21. Markovic D, Proll S, Bubenzer C, Scheer H. Myoglobin with chlorophyllous chromophores: influence on protein stability. Biochim Biophys Acta Bioenerg 2007;1767:897–904.
22. Proll S, Wilhelm B, Robert B, Scheer H. Myoglobin with modified tetrapyrrole chromophores: binding specificity and photochemistry. Biochim Biophys Acta Bioenerg 2006;1757:750–763.
23. Wright KA, Boxer SG. Solution properties of synthetic chlorophyllide- and bacteriochlorophyllide-apomyoglobin complexes. Biochemistry 1981;20:7546–7556.
24. Rau HK, Snigula H, Struck A, Robert B, Scheer H, Haehnel W. Design, synthesis and properties of synthetic chlorophyll proteins. Eur J Biochem 2001;268:3284–3295.
25. Braun P, Vegh AP, von Jan A, Strohmann B, Hunter CN, Robert B, Scheer H. Identification of intramembrane hydrogen bonding between 13(1) keto group of bacteriochlorophyll and serine residue α 27 in the LH2 light-harvesting complex. Biochim Biophys Acta Bioenerg 2003;1607:19–26.
26. Braun P, Fiedor L. Design and assembly of functional light-harvesting complexes. In: Hunter CN, Daldal F, Thurnauer MC, Beatty JT, editors. The purple phototrophic bacteria. Dorderecht: Springer; 2008. pp913–940.
27. Wang GL, Dunbrack RL. PISCES: a protein sequence culling server. Bioinformatics 2003;19:1589–1591.
28. Christopher JA, Baldwin TO. SPOCK: real-time collaborative molecular modelling. J Mol Graph Model 1998;16:285–285.
29. Kahn PC. Defining the axis of a helix. Comput Chem 1989;13:185–189.
30. Kahn PC. Simple methods for computing the least-squares line in three dimensions. Comput Chem 1989;13:191–195.
31. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
32. Summa CM. Computational methods and their applications for de novo functional protein design and membrane protein solubilization [Doctoral Thesis]. Philadelphia: University of Pennsylvania School of Medicine; 2002.
33. Bower MJ, Cohen FE, Dunbrack RL Jr. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. J Mol Biol 1997;267:1268–1282.
34. Chakrabarti P. Geometry of interaction of metal ions with histidine residues in protein structures. Protein Eng 1990;4:57–63.
35. Fufezan C, Zhang J, Gunner MR. Ligand preference and orientation in b- and c-type heme-binding proteins. Proteins: Struct Funct Bioinf 2008;73:690–704.
36. Mokrab Y, Stevens TJ, Mizuguchi K. Lipophobicity and the residue environments of the transmembrane α-helical bundle. Proteins: Struct Funct Bioinf 2009;74:32–49.
37. Silber MV, Gabriel G, Strohmann B, Garcia-Martin A, Robert B, Braun P. Fine tuning of the spectral properties of LH2 by single amino acid residues. Photosynth Res 2008;96:145–151.

38. Brust T, Draxler S, Rauh A, Silber MV, Braun P, Zinth WG, Braun M. Mutations of the peripheral antenna complex LH2-correlations of energy transfer time with other functional properties. Chem Phys 2009;357:28–35.
39. Gordon DB, Marshall SA, Mayo SL. Energy functions for protein design. Curr Opin Struct Biol 1999;9:509–513.
40. Balaban TS, Braun P, Hattig C, Hellweg A, Kern J, Saenger W, Zouni A. Preferential pathways for light-trapping involving beta-ligated chlorophylls. Biochim Biophys Acta Bioenerg 2009;1787:1254–1265.
41. Ho SP, DeGrado WF. Design of a 4-helix bundle protein: synthesis of peptides which self-associate into a helical protein. J Am Chem Soc 1987;109:6751–6758.
42. Das R, Baker D. Macromolecular modeling with rosetta. Annu Rev Biochem 2008;77:363–382.
43. Fry HC, Lehmann A, Saven JG, DeGrado WF, Therien MJ. Computational design and elaboration of a de novo heterotetrameric α-helical protein that selectively binds an emissive abiological (Porphinato)zinc chromophore. J Am Chem Soc 2010;132:3997–4005.