Profile Hidden Markov Models for Analyzing Similarities and Dissimilarities in the Bacterial Reaction Center and Photosystem II[†]

Eva-Maria Krammer,[‡] Pierre Sebban,[§] and G. Matthias Ullmann*,[‡]

Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstrasse 30, BGI, 95447 Bayreuth, Germany, and Laboratoire de Chimie Physique, UMR 8000, University P. XI/CNRS, Bât. 350, Faculté d'Orsay, 91405 Orsay Cedex, France

Received October 31, 2008; Revised Manuscript Received December 23, 2008

ABSTRACT: The bacterial photosynthetic reaction center is the evolutionary ancestor of the Photosystem II reaction center. These proteins share the same fold and perform the same biological function. Nevertheless, the details of their molecular reaction mechanism differ. It is of significant biological and biochemical interest to determine which functional characteristics are conserved at the level of the protein sequences. Since the level of sequence identity between the bacterial photosynthetic reaction center and Photosystem II is low, a progressive multiple-sequence alignment leads to errors in identifying the conserved residues. In such a situation, profile hidden Markov models (pHMM) can be used to obtain reliable multiple-sequence alignments. We therefore constructed the pHMM with the help of a sequence alignment based on a structural superposition of both proteins. To validate the multiple-sequence alignments obtained with the pHMM, the conservation of residues with known functional importance was examined. Having confirmed the correctness of the multiple-sequence alignments, we analyzed the conservation of residues involved in hydrogen bonding and redox potential tuning of the cofactors. Our analysis reveals similarities and dissimilarities between the bacterial photosynthetic reaction center and Photosystem II at the protein sequence level, hinting at different charge separation and charge transfer mechanisms. The conservation analysis that we perform in this paper can be considered as a model for analyzing the conservation in proteins with a low level of sequence identity.

A wide variety of species use photosynthesis to convert solar energy into chemical energy. Photosynthetic reaction centers (RCs)[1] of bacteria and eukaryotes perform a crucial step in this reaction, the conversion of excitation energy into redox energy (*1*−*3*). All RC proteins are evolutionarily related and share a common design (*1, 4*−*11*), which is characterized by a pseudosymmetric arrangement of the polypeptide chains and the bound cofactors. Nevertheless, RC proteins can be divided into two groups depending on the nature of the terminal electron acceptor, which is a Fe_4S_4 iron−sulfur cluster for type I RC proteins and a quinone (Q_B) for type II RC proteins. RC proteins from purple bacteria (bRC) are considered to be the ancestors of type II RC proteins of cyanobacteria, algae, and higher plants. The biological task of type II RCs is the utilization of absorbed light energy to transfer electrons to a lipid soluble quinone.

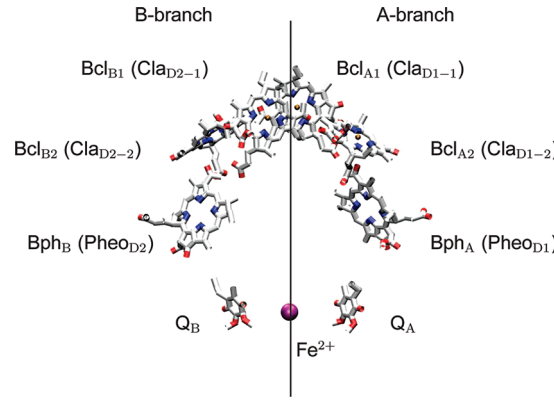


FIGURE 1: Overall organization of the cofactors in RC proteins. The cofactor molecules of the bRC of *Rhodobacter sphaeroides* are shown as an example. If the cofactors are named differently in bRC and PSII, the name in PSII is given in parentheses. The pseudo-symmetry axis of the cofactor molecules is schematically depicted. The figure was created using vmd (*80*) and is based on crystal structure 2C8J (*49*).

Figure 1 shows the typical orientation of the pigment and cofactor molecules in two symmetrically arranged branches (A and B) of the type II RC proteins. Only one branch, the A branch, is electron transfer active. Each branch consists of two (bacterio) chlorophyll molecules (Bcl or Chla), one (bacterio) pheophytin molecule (Bph or Pheo), and one quinone molecule (Q). To differentiate between the pigments in the two branches, the subscripts A and B are used. A non-heme iron ion is placed symmetrically between the two quinone molecules. The pigment and cofactor molecules are

[†] This work was supported by DFG Grant UL174/7-1 and by the DAAD German/French Procope Bilateral Travel Grant 11438QH and D/0502198.

* To whom correspondence should be addressed: Structural Biology/Bioinformatics, University of Bayreuth, Universitätsstr. 30, BGI, 95447 Bayreuth, Germany. Telephone: +49-921-55-3545. Fax: +49-921-55-3071. E-mail: matthias.ullmann@uni-bayreuth.de.

[‡] University of Bayreuth.

[§] University P. XI/CNRS.

[1] Abbreviations: Bcl, bacteriochlorophyll; Bph, bacteriopheophytin; bRC, bacterial reaction center; Chla, chlorophyll; MSA, multiple-sequence alignment; Pheo, pheophytin; PA, profile alignment; PDB, Protein Data Bank; pHMM, profile hidden Markov models; PSII RC, Photosystem II reaction center; RC, reaction center; SSA, structural sequence alignment; SSA+, SSA with additional sequences.

10.1021/bi802033k CCC: $40.75
Published on Web 01/21/2009

noncovalently bound to two membrane-spanning subunits, which are called L and M in bRC and D1 and D2 in Photosystem II (PSII), respectively. Subunits L and M and subunits D1 and D2 are evolutionarily related, share the same fold (five α-helices), and are structurally superimposable. Despite this large structural similarity, subunits L, M, D1, and D2 only share a level of sequence identity of ∼5% (*12*), which is a typical value for comparing random protein sequences (*13*). Apart from subunits L and M, the bRC protein can contain a H subunit and a C subunit in addition (*14*). PSII contains in addition to subunits D1 and D2 more than 20 subunits (*15*). Even if bRC and PSII RC both perform the same biological function, the mechanistic details of the charge separation differ. For example, the charge separation is initiated in the bRC at the chlorophyll dimer, the so-called special pair, while the charge separation in the PSII RC occurs most likely at the chlorophyll Cla_{D1-2} in the A branch (*16*−*20*). Also, the mechanism of photoprotection, which prevents radiation damage of the protein, is different in bRC and PSII RC proteins (*21*−*25*). Given the evolutionary relation between RC proteins, it is of great interest to know which functionally important residues are conserved throughout the protein family and which residues are replaced in the different subfamilies. Residues for which a strict conservation is expected are those directly involved in chlorophyll, pheophytin, iron, and quinone binding.

The environment of the pigments causes different electron transfer properties for the A branch and the B branch (*26*−*28*), which allows electron transfer only through the A branch. Also, the two quinone binding sites, Q_A and Q_B, function differently (*29*−*31*). Q_A is only transiently reduced by one electron and is never protonated in the reaction cycle, whereas Q_B accepts two electrons and two protons. This functional difference is caused by the different polarities of the binding pockets. In the bRC, two acidic residues (L212 and L213) are located in the Q_B site, whereas two unpolar residues can be found in the equivalent positions in the Q_A site (AlaM248 and AlaM249). It is likely that these two residues (L212/L213 and M248/M249) largely determine the different redox character of the quinones in the sites of Q_A and Q_B in the bRC (*31, 32*).

The aim of this work is to analyze whether functional differences and similarities of bRC and PSII RC can be identified at the level of their sequences. A common strategy is to use a multiple-sequence alignment (MSA) for this analysis. A straightforward way to generate MSAs is by using progressive alignment algorithms as implemented for instance in ClustalW (*33*). However, MSAs obtained by a progressive alignment algorithm might not be reliable, especially when aligning sequences of proteins with a low level of sequence identity (i.e., with <20−30%) (*34, 35*). The use of structural information in the generation of MSAs makes them more reliable, since more information is used for these alignments (*36*−*40*). A reliable MSA can be used to construct a profile hidden Markov model (pHMM), which is an elegant way of representing the information contained in the MSA (*41*−*47*). Once a pHMM is generated, additional sequences can be aligned to the pHMM. This strategy gives reliable MSAs when the MSA used to generate the pHMM was sufficiently large and correct (*41*).

To differentiate between different kinds of alignments, we use the following definitions. (1) The term structural sequence alignment (SSA) stands for a MSA in which structural information is included into the alignment process. For this alignment, we use only sequences for which a crystal structure is available. (2) The term structure-based MSA (SSA+) stands for a MSA in which we align additional sequences with a SSA. (3) A profile alignment (PA) is a MSA obtained from aligning sequences to a pHMM. The residue numbering used in this paper is the numbering of the purple bacterium *Rb. sphaeroides*. For residues that are functionally important in PSII RC proteins only, we used the residue numbering of the thermophilic cyanobacterium *Thermosynechococcus elongatus*.

In this paper, we constructed separately two pHMMs, one for the L/D1 subunit and one for the M/D2 subunit. The pHMMs are based on SSA+s. Using these pHMMs, we obtained PAs which we used to examine whether functional differences and similarities between bRC and PSII RC are manifested on the sequence level. Such a strategy reduces the danger of obtaining and analyzing misalignments that can be found using progressive alignment algorithms. The PAs confirmed that the residues involved in the binding of the pigments and the cofactors are conserved in the RC proteins. The two quinone binding sites are similar in the RC proteins, but interesting functional differences can be found between the bRC and PSII RC. The properties of the residues surrounding the electron transfer sites in the A branch and the B branch are compared. While the hydrogen bonding patterns of the chlorophyll and pheophytin are conserved, many residues directing the electron transfer along the A branch in the bRC are not conserved in PSII RC proteins. This finding suggests a different mechanism for directing the electron transfer along the A branch in PSII. Thus, even though the bRC and PSII RC proteins are structurally similar and perform the same biological reaction, the molecular details of this reaction differ between the bRC and PSII RC.

MATERIALS AND METHODS

To obtain the SSAs and the SSA+s, the program staccato of the BioInfo3D software package (*36*) was used. The consideration of structural information in the generation of SSAs makes these alignments more reliable than progressive MSAs (*40, 48*). The structural information needed for the SSA was taken from the bRC and PSII RC crystal structures from *Rb. sphaeroides* (bRC, PDB entry 2J8C) (*49*), from *Blastochloris viridis* (bRC, PDB entry 1PRC) (*50*), and *T. elongatus* (PSII, PDB entry 2AXT) (*15*). We obtained two SSAs, one for the L/D1 subunit and one for the M/D2 subunit. With these SSAs, we aligned 47 additional homologous sequences to obtain the SSA+s using the program staccato. These homologous sequences were searched using the BLAST (*51*) algorithm as implemented on the NCBI webpage (*52*) starting from the sequences of the bRC from *Rb. sphaeroides* and of the PSII RC from *T. elongatus*. To obtain the SSA+ for the L/D1 subunit, we first aligned all considered L sequences and all considered D1 sequences separately with the SSA and subsequently merged these alignments to one alignment. To obtain a correct alignment of the functional key residue SerL223, the merged alignments needed to be adjusted by shifting the gap before SerL223 to the position after L223 (as it was found in the SSA) in the

Table 1: List of Marker Residues and Associated Sequence Words[a]

subunit	residue (bRC)	residue (PSII)	sequence word	function in the protein	symmetry (bRC)	symmetry (PSII)
L/D1	HisL153	*ThrD1-179*	WTHLD	ligand of Bcl_{A2}	M182	*IleD2-178*
	HisL173	D1-198	PAHMI	ligand of Bcl_{A1}	M202	D2-197
	HisL190	D1-215	ALHGA	coordination of the non-heme iron	M219	D2-214
	GluL212	AlaD1-251	DHEDTF	Q_B site	AlaM248	AlaD2-249
	AspL213	HisD1-252	DHEDTF	Q_B site	AlaM249	*AsnD2-250*
	SerL223	D1-264	GYSIG	Q_B site	*AsnM259*	*AlaD2-260*
	HisL230	D1-272	GIHRL	coordination of the non-heme iron	M266	D2-268
M/D2	HisM182	*IleD2-178*	FSHLD	ligand of Bcl_{B2}	L153	*ThrD1-179*
	HisM202	D2-197	PFHGL	ligand of Bcl_{B1}	L173	D1-198
	HisM219	D2-214	AMHGA	coordination of the non-heme iron	L190	D1-215
	GluM234	extra loop	ERELE	coordination of the non-heme iron	extra loop	extra loop
	AlaM248	D2-249	ERAALF	Q_A site	GluL212	D1-251
	AlaM249	*AsnD2-250*	ERAALF	Q_A site	AspL213	HisD1-252
	HisM266	D2-268	GIHRW	coordination of the non-heme iron	M230	D1-272
others	PheL216	D1-255	TFFRD	stacking to the quinone	TrpM252	TrpD2-253
	CysL92	HisD1-118[b]	ALHGA[b]	ligand of axial Chl (Chl_Z in PSII	*PheM121*	D2-117
	ArgL135	TyrD1-162[b]	LIYPL[b]	tyrosine radical (water cleavage) in PSII	*ArgM164*	D2-164

[a] Marker residues are residues that are known to be conserved in the considered proteins. A sequence word consists of the marker residue and two residues before and after the marker residue. The sequence words refer to the bRC of *Rb. sphaeroides* if not stated otherwise. In case a certain residue is known to be conserved in only a certain subunit, in bRC, or in PSII, the corresponding residues of the other subunits, for which no conservation information exists, are shown in italics. GluM234 is taken as a marker residue only for the M/D2 subunit. [b] No conservation in bRC (sequence word refers to PSII RC of *T. elongatus*).

sequences of the bRC proteins. For the alignment of the M/D2 subunit, we proceeded analogously. The so-constructed SSA+s were used to build pHMM for the L/D1 subunit and for the M/D2 subunit using the program hmmbuild of the software package HMMER (*43*).

To validate the obtained SSA+s and the PAs, we defined residues, named marker residues in the following, which are known to be conserved in the considered proteins. The marker residues are listed in Table 1. Most of them are involved in cofactor binding. Some marker residues are specific for only one subunit (L/D1 or M/D2), for only the bRC, or for only the PSII RC. For the validation of the alignments, we counted misalignments of each marker residue. For this purpose, we defined sequence words around each marker residue that consist of two residues before and two residues after the marker residue. A misalignment was identified if a sequence word that is in three of five residues identical to the defined marker residue sequence word occurs shifted in the alignment. The use of sequence words to identify misalignments is required, since in some of the sequences, a different amino acid can be found at a marker residue position. If the marker residues are aligned properly, one can assume that also all other residues are aligned correctly as well.

For the validation of the pHMM and for the conservation analysis of functional key residues, PAs using a large variety of sequences of subunits L, M, D1, and D2 were performed. These sequences were taken from a BLAST (*51*) search of the NCBI webpage (*52*). We carefully examined the sequences and removed multiple entries of the same species as well as mutated sequences. The deletion of such sequences is necessary, since they could otherwise bias the conservation analysis. We considered in total 616 sequences: 100, 114, 165, and 237 sequences for subunits L, M, D1, and D2, respectively. Two PAs of these 616 sequences were generated, one using the pHMM of the L/D1 subunit and another one using the pHMM of the M/D2 subunit. Since some of the sequences for subunits L and M are not full-length sequences, not all sequences could be used for the validation of all marker residues and for further analysis. Complete lists of the sequences used for building, validating, and analyzing the two obtained pHMMs, the obtained alignments, and the pHMM files are given in the Supporting Information.

RESULTS AND DISCUSSION

Failure of Progressive Alignment and an Existing pHMM. As a first approach to obtaining a MSA of the type II RC sequences, we performed a progressive alignment using ClustalW (*33*). These calculations did not lead to correct alignments of the marker residue SerL223 using various sequences for subunits L and D1. For the marker residues GluL212 and AspL213 of the bRC, which are Ala and His, respectively, in the PSII RC D1 subunit, a correct MSA was only obtained for some sets of sequences, but not for all. Changing the scoring matrix did not improve the results. A figure showing examples of the misaligned sequences is available as Supporting Information.

The Pfam database (*53, 54*) contains a single pHMM file for subunits L, M, D1, and D2 together. The Pfam pHMM succeeds in aligning residues responsible for chlorophyll and non-heme iron ligation but fails to align important residues of the Q_A and Q_B sites. Thus, the pHMM of Pfam may succeed in classifying the considered sequences as type II RC sequences, which is the actual purpose of the Pfam database. However, the Pfam pHMM does not align all functionally relevant residues reliably.

Performance of the New pHMMs. To prevent the problem of misaligning important residues, we generated two separate pHMMs from SSA+s, one for the L/D1 subunit and the other for the M/D2 subunit. The SSA led to a proper alignment of all marker residues (see Figure 2). Table 2 shows the analysis of the PA of the 616 sequences using either the L/D1 pHMM or the M/D2 pHMM. Both PAs succeed in aligning all histidine marker residues L153, L173, L190, and L230 and M183, M202, M219, and M266 which are involved in metal coordination. Since the Bcl- and iron-coordinating histidines are present in the L/D1 subunit and also in the M/D2 subunit, this result is expected. In addition for PSII RC proteins, the

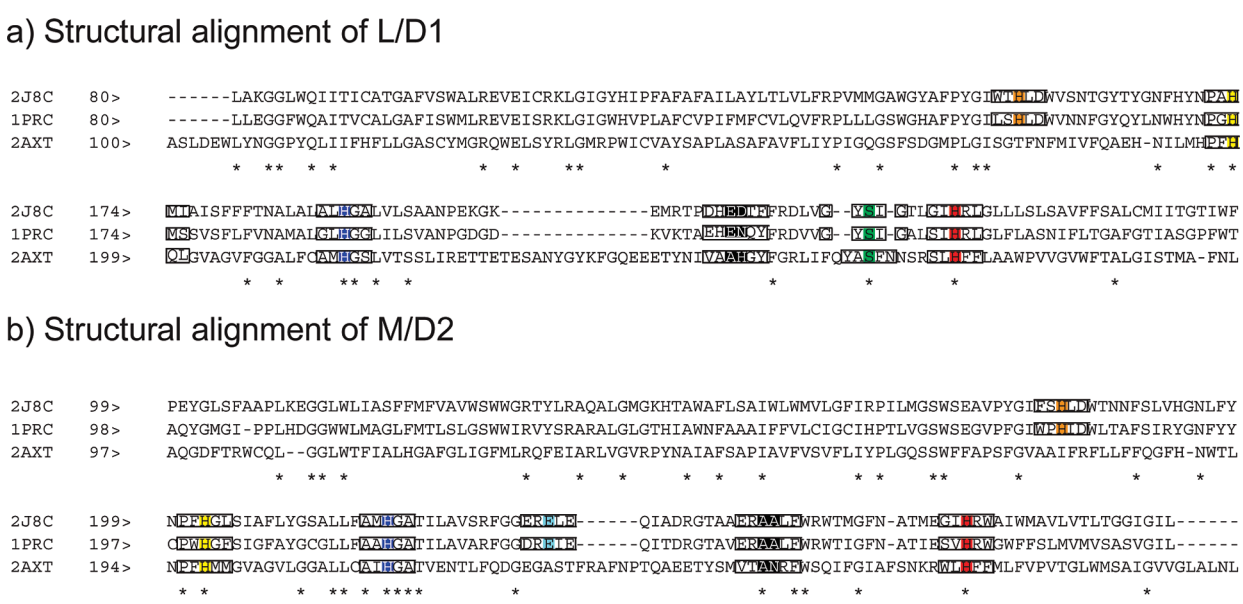


FIGURE 2: Part of the SSA of the L/D1 subunit (a) and the M/D2 subunit (b). Marker residues and their corresponding sequence words are highlighted. The sequences of the bRC proteins of *Rb. sphaeroides* (PDB entry 2J8C) (*49*) and *B. viridis* (PDB entry 1PRC) (*50*) and of the PSII RC of *T. elongatus* (PDB entry 2AXT) (*15*) were used for the SSA. (a) The marker residues HisL153 (orange), HisL173 (yellow), HisL190 (blue), L212/L213 (black), SerL223 (green), HisL230 (red), HisM182 (orange), HisM202 (yellow), and HisM219 (blue) are depicted. (b) The marker residues HisM182 (orange), HisM202 (yellow), HisM219 (blue), GluM234 (cyan), M248/249 (black), and HisM266 (red) are depicted. In the sequence word, the marker residues are colored using the same color for equivalent residues, i.e., those that have the same position in L/M and D1/D2 subunits.

Table 2: Analysis of the Profile Alignment of 616 Sequences (100, 114, 165, and 237 sequences for L, M, D1, and D2, respectively) Using either the L/D1 pHMM or the M/D2 pHMM[a]

		no. of misaligned sequences per total number of sequences			
pHMM	residue	L subunit	M subunit	D1 subunit	D2 subunit
L/D1	HisL153	0/100	0/113	0/165	0/237
	HisL173	0/100	0/111	0/165	0/237
	HisL190	0/100	0/111	0/165	0/237
	GluL212	0/96	53/95 (55.8%)	0/165	0/237
	AspL213	0/95	53/95 (55.8%)	0/165	0/237
	SerL223	0/93	43/49 (87.8%)	0/165	0/237
	HisL230	0/95	0/49	0/165	0/237
	GluM234	–	22/49 (44.9%)	–	–
	PheL216	0/95	29/75 (38.7%)	0/165	0/237
	HisD1-118	0/100	0/112	0/165	0/234
	TyrD1-162	0/100	0/113	0/165	0/237
M/D2	HisM182	0/100	0/113	0/165	0/237
	HisM202	0/100	0/111	0/165	0/237
	HisM219	1/100	0/111	0/165	0/237
	GluM234	–	0/49	–	–
	AlaM248	38/96 (39.6%)	0/95	0/165	0/237
	AlaM249	38/95 (40.0%)	0/95	0/165	0/237
	HisM266	0/93	0/49	0/165	0/237
	SerL223	2/93 (2.2%)	0/49	165/165 (100%)	0/237
	PheL216	0/95	0/75	165/165 (100%)	0/237
	HisD1-118	0/100	0/113	0/165	0/234
	TyrD1-162	0/100	0/113	0/165	0/237
identity	lowest	57.5%	48.8%	80.8%	73.6%
identity	average	70.2%	64.3%	86.4%	89.7%

[a] The analysis shows that the L/D1 pHMM fails to align some marker residue in the M subunit. The same is observed when the pHMM of the M/D2 subunit is used to align sequences of the L subunit. Thus, separate pHMMs for the L/D1 and M/D2 subunits are needed to align all marker residues correctly. The percentages of misalignment are shown in parentheses. If a sequence is not used for the validation of a certain marker residue, since the sequence is too short, the total number of sequences is reduced. The sequence identity ranges of the analyzed sequences for subunits L, M, D1, and D2 with respect to the respective subunit of the bRC of *Rb. sphaeroides* (L and M) and to the sequence of the PSII RC of *T. elongatus* (D1 and D2) are listed in the last two rows.

functionally important residues HisD1-118, coordinating the axial chlorophyll Cla_Z, and TyrD1-162, involved in water splitting, are properly positioned in both PAs. A proper alignment of the marker residues situated in the Q_A and Q_B binding site is only achieved if the appropriate pHMM is used, i.e., the pHMM of the M/D2 subunits for the Q_A site and the pHMM of the L/D1 subunits for the Q_B site. A correct alignment of GluM234, the fifth ligand of the iron in the bRC, is only achieved when the pHMM of the M/D2 subunits is used (Table 2). Thus, two separate pHMMs for two subunits are definitely required to obtain a proper alignment of all functionally important residues.

Table 3: Conservation of Residues near the Oxygen Atoms of the Special Pair[a]

		bRC			PSII RC		
chlorophyll	atom name	residue	conservation (%)	exchanged to (%)	residue	conservation (%)	exchanged to (%)
Bcl_{A1}	O1A	SerL244	97.6	G (2.4)	ThrD1-286	99.4	A (0.6)
	O2A	AlaL127	98.0	M (2.0)	AlaD1-154	97.0	S/P (1.8/1.2)
		PheL181	98.0	W/L (1.0/1.0)	PheD1-206	99.4	L (0.6)
		ValL241	36.1	G/A/F (42.2/19.4/2.3)	ValD1-283	1.2	I (98.8)
		SerL244	97.6	G (2.4)	ThrD1-286	99.4	A (0.6)
	O1D	SerL244	97.6	G (2.4)	ThrD1-286	99.4	A (0.6)
		CysL247	97.6	G (2.4)	GlyD1-289	96.4	S (3.6)
		MetL248	15.7	I/T (81.9/2.4)	IleD1-290	76.3	V/L/N (20.6/2.4/0.6)
	O2D	LeuL131	100.0		ValD1-157	98.8	I (1.2)
		SerL244	97.6	G (2.4)	ThrD1-286	99.4	A (0.6)
		MetL248	15.7	I/T (81.9/2.4)	IleD1-290	76.3	V/L/N (20.6/2.4/0.6)
	OBD	LeuL131	100.0		ValD1-157	98.8	I (1.2)
		MetL248	15.7	I/T (81.9/2.4)	IleD1-290	76.3	V/L/N (20.6/2.4/0.6)
	OBB	HisL168	95.0	E (5.0)	LeuD1-193	99.4	Y (0.6)
		ThrM186	100.0		LeuD2-182	26.2	I (73.8)
Bcl_{B1}	O1A	PheL181	98.0	W/L (1.0/1.0)	PheD1-206	99.4	L (0.6)
		LeuM156	91.2	F/V (7.9/0.9)	ValD2-152	94.5	I/L (5.1/0.4)
		LeuM209	97.3	A/M (1.8/0.9)	ValD2-204	73.0	I (27.0)
		ValM276	21.3	T/C (55.3/23.4)	GlyD2-278	100.0	
		ThrM277	40.4	P/V/M/C (42.6/12.8/2.1/2.1)	LeuD2-279	92.8	M/S (6.8/0.4)
	O2A	PheL181	98.0	W/L (1.0/1.0)	PheD1-206	99.4	L (0.6)
		LeuM156	91.2	F/V (7.9/0.9)	ValD2-152	94.5	I/L (5.1/0.4)
		ThrM277	40.4	P/V/M/C (42.6/12.8/2.1/2.1)	LeuD2-279	92.8	M/S (6.8/0.4)
	O1D	LeuM156	91.2	F/V (7.9/0.9)	ValD2-152	94.5	I/L (5.1/0.4)
		SerM205	97.2	C/A (1.9/0.9)	GlyD2-200	100.0	
		GlyM280	87.3	A/S (10.6/2.1)	SerD2-282	98.4	A/C/N (0.8/0.4/0.4)
		IleM284	97.7	V (2.3)	ValD2-286	70.9	I/L (28.3/0.8)
	O2D	GlyM280	87.3	A/S (10.6/2.1)	SerD2-282	98.4	A/C/N (0.8/0.4/0.4)
		IleM284	97.7	V (2.3)	ValD2-286	70.9	I/L (28.3/0.8)
	OBD	LeuM156	91.2	F/V (7.9/0.9)	ValD2-152	94.5	I/L (5.1/0.4)
		LeuM160	92.0	I/V/C (6.2/0.9/0.9)	ValD2-156	99.6	C (0.4)
		TrpM185	100.0		PheD2-181	99.6	S (0.4)
		GlyM280	87.3	A/S (10.6/2.1)	SerD2-282	98.4	A/C/N (0.8/0.4/0.4)
	OBB	TyrM210	99.1	F (0.9)	LeuD2-205	99.6	P (0.4)
		AsnL166	34.0	H (60.0)	AsnD1-191	100.0	
		MetL174	99.0	C (1.0)	GlnD1-199	7.9	M/E/I (90.9/0.6/0.6)
		PheM197	55.0	Y (45.0)	ThrD2-192	100.0	

[a] Residues near oxygen atoms of the C_7 phytyl (O1A and O2A), C_{10} carbomethyl (O1D and O2D), C_9 keto (OBD), and C_{2a} (OBB) acetyl groups of the special pair are determined using crystal structure 2J8C (*49*). The names given in parentheses are the PDB atom names. The residue numbering of bRC and PSII RC refers to *Rb. sphaeroides* and *T. elongatus*, respectively.

Sequence-Based Comparison of the bRC and PSII RC. The PAs, i.e., the MSA obtained by using the pHMMs, are used to detect differences and similarities of bRC and PSII RC proteins. We focused on hydrogen bonds to the pheophytin and chlorophyll pigments, on redox tuning of the A and B branches, and on the properties of the Q_A site and the Q_B site. Residues which are near the special pair, the accessory chlorophyll, and the pheophytin pigment molecules and their conservation are listed in Tables 3, 4, and 5, respectively. The conservation of residues involved in redox potential tuning of the A branch and the B branch is shown in Table 6. The conservation of functionally important residues in the Q_A site and the Q_B site is listed in Table 7. A residue is considered to be conserved, if the same amino acid is found in at least 90% of all analyzed sequences of the same subunit (L, M, D1, and D2) and only replaced by amino acids of the same character in the remaining sequences. An amino acid exchange that is seen in only one sequence is considered to be a sequencing artifact. It should be kept in mind that conserved residues could be important for different reasons. They could be involved in cofactor binding, in folding and assembly, or in the interaction with other proteins or subunits. A table of highly conserved residues is given as Supporting Information. For further analysis of the RC sequences, we provide the PAs and the pHMM files as Supporting Information. The conservation analysis performed in this work is described in more detail in the following subsections.

(i) Hydrogen Bonding and Midpoint Potential Tuning of the Special Pair. Since it was shown experimentally that the midpoint potential of the special pair can be altered by removing or adding a hydrogen bond to the special pair (*55*−*61*), we examined which possible hydrogen bond partners exist and if the character of these amino acids is conserved. A bacteriochlorophyll molecule contains six oxygen atoms (see Figure 3) which are possible hydrogen bond acceptors: the C_9 keto oxygen, the C_{2a} acetyl oxygen, the two C_{10} methyl ester oxygen atoms, and the two oxygen atoms of the ester in the C_7 phytyl chain. We determined all residues that are at least 5 Å from the oxygen atoms of the special pair in the bRC of *Rb. sphaeroides* (PDB entry 2J8C). This distance was chosen to be larger than the normal length of a hydrogen bond (3.6 Å) to take structural variability into account. In Table 3, the identified residues close to the oxygen atoms of the special pair and their conservation are listed. Their location in the bRC structure is depicted in Figure 4.

The C_7 phytyl oxygen atoms of Bcl_{A1} can form a hydrogen bond with the side chain of SerL244 in the bRC of *Rb. sphaeroides*, which is found in 97.6% of the bRC sequences. In the few cases when SerL244 is replaced, it is only

Table 4: Conservation of Residues near the Oxygen Atoms of the Accessory Chlorophyll Molecules[a]

		bRC			PSII RC		
chlorophyll	atom name	residue	conservation (%)	exchanged to (%)	residue	conservation (%)	exchanged to (%)
Bcl_{A2}	O1A	AlaM207	42.3	V/F/G (51.4/4.5/1.8)	AlaD2-202	99.6	R (0.4)
		GlyM211	100.0		GlyD2-206	100.0	
	O2A	AlaM207	42.3	V/F/G (51.4/4.5/1.8)	AlaD2-202	99.6	R (0.4)
		TyrM210	99.1	F (0.9)	LeuD2-205	99.6	P (0.4)
		GlyM211	100.0		GlyD2-206	100.0	
	O1D	LeuL154	100.0		PheD1-180	100.0	
		GlyM203	7.2	A/M/C/V (42.3/28.8/19.8/1.8)	MetD2-198	99.6	T (0.4)
	O2D	IleL150	97.0	T/V/F (1.0/1.0/1.0)	IleD1-176	100.0	
		TrpL151	17.0	F/M/L/I/Y (44.0/17.0/12.0/7.0/3.0)	SerD1-177	99.4	T (0.6)
		LeuL154	100.0		PheD1-180	100.0	
	OBD	LeuL154	100.0		PheD1-180	100.0	
		GlyM203	7.2	A/M/C/V (42.3/28.8/19.8/1.8)	MetD2-198	99.6	T (0.4)
		IleM206	100.0		ValD2-201	99.6	I (0.4)
	OBB	PheL97	100.0		AlaD1-123	0.6	V/I/C (75.8/21.8/1.8)
		TyrL128	94.0	F (6.0)	AlaD1-154	97.6	T/S (1.8/0.6)
		PheL146	100.0		MetD1-172	98.2	I/L (1.2/0.6)
Bcl_{B2}	O1A	ThrL182	87.0	A/G/I/T/V (8.0/2.0/1.0/1.0/1.0)	GlyD1-207	100.0	
	O2A	SerL178	28.0	T/A/V (68.0/3.0/1.0)	AlaD1-203	100.0	
		PheL181	98.0	W/L (1.0/1.0)	PheD1-206	99.4	L (0.6)
		ThrL182	87.0	A/G/I/T/V (8.0/2.0/1.0/1.0/1.0)	GlyD1-207	100.0	
	O1D	MetL174	99.0	C (1.0)	GlnD1-199	7.9	M/E/I (90.9/0.6/0.6)
		SerL178	28.0	T/A/V (68.0/3.0/1.0)	AlaD1-203	100.0	
	O2D	MetL174	99.0	C (1.0)	GlnD1-199	7.9	M/E/I (90.9/0.6/0.6)
		LeuM183	98.2	P/I (0.9/0.9)	PheD2-179	99.2	S (0.8)
	OBD	HisL168	95.0	E (5.0)	LeuD1-193	99.4	Y (0.6)
		MetL174	99.0	C (1.0)	GlnD1-199	7.9	M/E/I (90.9/0.6/0.6)
		IleL177	19.0	V/L (80.0/1.0)	ValD1-202	100.0	
	OBB	MetM122	6.2	L/F (85.8/8.0)	GlyD2-122	99.6	A (0.4)
		TrpM157	3.5	F/Y/V/M/I (48.7/36.3/8.0/2.7/0.9)	PheD2-182	99.2	S (0.8)

[a] Residues near oxygen atoms of the C_7 phytyl (O1A and O2A), C_{10} carbomethyl (O1D and O2D), C_9 keto (OBD), and C_{2a} (OBB) acetyl groups of the accessory chlorophyll molecules are determined using crystal structure 2J8C (*49*). The names given in parentheses are the PDB atom names. The residue numbering of bRC and PSII RC refers to *Rb. sphaeroides* and *T. elongatus*, respectively.

exchanged with the smaller amino acids, Ala and Gly. A water molecule that could bind in the position of the hydroxyl group may then form a hydrogen bond with the C_7 phytyl oxygen atoms of Bcl_{A1}. In the B branch of the bRC, ValM276 and ThrM277 are close to the C_7 phytyl oxygen atoms of Bcl_{B1}. However, the orientation of ThrM277 does not allow the formation of a hydrogen bond in the considered structure. Moreover, this residue is exchanged in 57.5% of the sequences with amino acids, which are not able to form hydrogen bonds. In most of the sequences, ValM276 is replaced with Thr or Cys which can in principle form a hydrogen bond. All residues near the C_7 phytyl chain oxygen atoms in bRC are conserved as amino acids with the same hydrogen bonding properties in the PSII RC (see Table 3), except ValM276 which is a Gly in PSII. However, in the structure of the PSII RC, it seems possible that a cavity large enough to hold a water molecule is located near the C_7 phytyl chain oxygen atoms of $Cla_{D2\text{-}1}$. Thus, in analogy to the bRC, the C_7 phytyl chain oxygen atoms of $Cla_{D1\text{-}1}$ of PSII are singly hydrogen bonded to ThrD1-286, and a hydrogen bond with the C_7 phytyl chain oxygen atoms of $Cla_{D2\text{-}1}$ is possible.

The C_{10} carboxymethyl ester oxygen atoms of Bcl_{A1} can form a hydrogen bond with SerL244 and CysL247 in the bRC of *Rb. sphaeroides*. SerL244 and CysL247 are only exchanged with smaller amino acids (Ala and Gly) in the analyzed bRC sequences, which allows that a water binds in the position of the hydroxyl or thiol group. SerL244 is between the C_7 phytyl oxygen atom and C_{10} carboxymethyl ester oxygen atom O1D of Bcl_{A2} and may thus form a hydrogen bond with either of the two oxygen atoms. In PSII RC proteins, ThrD1-286 is between two C_{10} carboxymethyl ester oxygen atoms of $Cla_{D1\text{-}1}$ and can possibly form a hydrogen bond with either of them, resembling the situation of SerL244 in the bRC. In the B branch of the bRC, SerM205, GlyM280, and IleM284 are all close to the C_{10} carboxymethyl ester oxygen atoms of Bcl_{B1}. Interestingly among them, there is always one amino acid that can form a hydrogen bond. Of these three residues, only one, SerD2-282 which is equivalent to GlyM280 in the bRC, can form a hydrogen bond with the C_{10} carboxymethyl ester oxygen atoms of Bcl_{B1} in the PSII RC. The hydrogen bonding to the C_{10} carboxymethyl oxygen and the C_7 phytyl oxygen atoms of the special pair are important for the proper positioning of the chlorophyll molecules in type II RC proteins. A strict conservation of the hydrogen bonding partners seems therefore less important.

In contrast to the above-discussed oxygen atoms, the C_9 keto and C_{2a} acetyl oxygen atoms are part of the conjugated π-system of chlorophyll molecules. Hydrogen bonding to these atoms therefore influences the redox potential. Apart from 2.4% of the sequences, in which a Thr is found at position L248, no hydrogen bond is formed to the C_9 keto oxygen of Bcl_{A1} in bRC proteins. Interestingly, one of the bRC proteins with a Thr at position L248 is the bRC of *B. viridis*, for which crystal structures are available (*50*). Near the C_9 keto oxygen atom of Bcl_{B1}, only conserved hydrophobic residues (see Table 3) are located in the bRC, and thus, no hydrogen bond is formed. Also in PSII, no hydrogen bond is formed to the C_9 keto oxygen atoms of the special pair. Investigations using the bRC of *Rb. sphaeroides* showed that the mutation of LeuL131 or LeuM160, which introduces a hydrogen bond to the C_9 keto oxygen atoms of the special

Table 5: Conservation of Residues near the Oxygen Atoms of the Pheophytin Molecules[a]

		bRC			PSII RC		
pheophytin	atom name	residue	conservation (%)	exchanged to (%)	residue	conservation (%)	exchanged to (%)
Bph_A	O1A	AlaL96[b]	88.9	S/V/T/R (7.1/1.0/1.0/1.0)	GlyD1-122	99.4	S (0.6)
		PheL97	100.0		AlaD1-123	0.6	V/T/C (75.8/21.8/1.8)
		TrpL100[b]	99.0		TyrD1-126	96.4	W (3.6)
	O2A	ThrL38	71.6	A/G/V/S/E (14.8/6.8/3.4/2.3/1.1)	MetD1-37	99.4	L (0.6)
		PheL97	100.0		AlaD1-123	0.6	V/I/C (75.8/21.8/1.8)
	O1D	TrpL100[b]	99.0		TyrD1-126	96.4	W (3.6)
		MetM218	97.2	A/R (1.9/0.9)	IleD2-213	100.0	
		MetM256	98.6	I (1.4)	PheD2-257	99.2	S/L (0.4/0.4)
	O2D	TrpL100[b]	99.0		TyrD1-126	96.4	W (3.6)
		GluL104	72.0	Q (28.0)	GlnD1-130	9.7	E (90.3)
		ThrM255	97.2	Q/C (1.4/1.4)	IleD2-256	97.0	V/L/T (2.2/0.4/0.4)
		MetM256	97.2	I/L (1.4/1.4)	PheD2-257	99.2	S/L (0.4/0.4)
	OBD	TrpL100[b]	99.0		TyrD1-126	96.4	W (3.6)
		GluL104	72.0	Q (28.0)	GlnD1-130	9.7	E (90.3)
		IleL117	19.0	V/L (80.0/1.0)	IleD1-143	100.0	
	OBB	TyrM210	99.1	F (0.9)	LeuD2-205	99.6	P (0.4)
Bph_B	O1A	AlaM125	14.1	S/G (85.0/0.9)	GlyD2-121	75.5	A (24.5)
		ValM126	46.9	I/L/C (37.2/15.0/0.9)	LeuD2-122	94.9	V/I/S (3.8/0.9/0.4)
		TrpM129	99.1	F (0.9)	PheD2-125	99.6	P (0.4)
		PheM150	99.1	V (0.9)	PheD2-146	100.0	
	O2A	AlaM125	14.1	S/G (85.0/0.9)	GlyD2-121	75.5	A (24.5)
		TrpM129	99.1	F (0.9)	PheD2-125	99.6	P (0.4)
	O1D	LeuL219	34.0	F/T/I/V/D/A/M (31.9/11.7/8.5/7.4/3.2/2.1/1.1)	LeuD1-258	100.0	
		ValL220	26.6	I/L/M (70.2/2.1/1.1)	IleD1-259	89.1	V/F/A/G/T (3.6/2.4/1.2/1.8/1.8)
		TrpM129	99.1	F (0.9)	PheD2-125	99.6	P (0.4)
		ThrM133	64.7	V/S/M/A/I (10.6/8.8/8.8/3.5/3.5)	GlnD2-129	100.0	
	O2D	LeuL185	78.0	F/M/A/W (15.0/5.0/2.0)	LeuD1-212	100.0	
		LeuL189	80.0	M (20.0)	MetD1-214	100.0	
		LeuL219	34.0	F/T/I/V/D/A/M (31.9/11.7/8.5/7.4/3.2/2.1/1.1)	LeuD1-258	100.0	
	OBD	LeuL189	80.0	M (20.0)	MetD1-214	100.0	
		LeuL219	34.0	F/T/I/V/D/A/M (31.9/11.7/8.5/7.4/3.2/2.1/1.1)	LeuD1-258	100.0	
		ThrM133	64.7	V/S/M/A/I (10.6/8.8/8.8/3.5/3.5)	GlnD2-129	100.0	
		ThrM146	14.2	V/I/L/M/R (42.5/30.1/7.1/5.3/0.9)	AsnD2-142	100.0	
		PheM150	99.1	V (0.9)	PheD2-146	100.0	
	OBB	PheL181	98.0	W/L (1.0/1.0)	PheD1-206	99.4	L (0.6)

[a] Residues near the oxygen atoms of the C_7 phytyl (O1A and O2A), C_{10} carbomethyl (O1D and O2D), C_9 keto (OBD), and C_{2a} (OBB) acetyl groups of the pheophytin molecules are determined using crystal structure 2J8C (*49*). The names given in parentheses are the PDB atom names. The residue numbering of bRC and PSII RC refers to *Rb. sphaeroides* and *T. elongatus*, respectively. [b] In one bRC sequences, a gap is at this position.

pair, alters the redox potential of the special pair (*56, 60, 61*). Considering the conservation of the residues near the C_9 keto oxygen atoms, the redox potential of the special pair is not altered by hydrogen bonds to these atoms in wild-type RC proteins. The C_{2a} acetyl group is present in only bacteriochlorophyll and not in chlorophyll *a* of PSII (see Figure 3). In the bRC, the conserved residues HisL168 and ThrM186 are close to the C_{2a} acetyl oxygen atom of Bcl_{A1}. However, ThrM186 forms a hydrogen bond with the backbone carbonyl of HisM182, and therefore, its orientation does not allow hydrogen bonding to Bcl_{A1}. Thus, the C_{2a} acetyl oxygen atom of Bcl_{A1} accepts only one hydrogen bond which originates from HisL168. The removal of this hydrogen bond (by mutation) alters the redox potential of the special pair in *Rb. sphaeroides* (*56, 57, 60*). The strength of this hydrogen bond is influenced by AsnL166. Mutation of AsnL166 leads to an alteration of the redox potential of the special pair in the bRC of *Rb. sphaeroides* (*57*). Our analysis shows that AsnL166 is not strictly conserved, but it is always exchanged with an amino acid, which can form a hydrogen bond. In the B branch, only the conserved TyrM210 is located near the C_{2a} acetyl oxygen of Bcl_{B1} in the bRC. Although some experiments suggested that this residue does not form a hydrogen bond with the C_{2a} acetyl oxygen of Bcl_{B1} (*62*), the structural arrangement makes a hydrogen bond plausible.

It has been shown that in the bRC of *B. viridis* and *Chromatium tepidum* a tyrosine at position M197 forms a hydrogen bond to the C_{2a} acetyl oxygen of Bcl_{B1} (*50, 63*). In the bRC of *Rb. sphaeroides*, the PheM197 → Tyr mutation leads to a new hydrogen bond and increases the redox potential of the special pair (*55, 64*). At position M197, both a tyrosine and a phenylalanine are found in the bRC sequences, seemingly implying different redox potentials of the special pair in the respective bRC proteins. However, the redox potentials of the special pair in the bRC of *Rb. sphaeroides* and *B. viridis* are similar (~500 mV) (*56, 65*). A clear difference between the bRC of *Rb. sphaeroides* and the bRC of *B. viridis* is the presence of a C subunit in the latter species. The C subunit shields the special pair more strongly from the solvent, which would cause a shift of its redox potential (*66*). The additional hydrogen bond in the bRC of *B. viridis* can compensate for this effect. We therefore analyzed whether the presence of a C subunit is correlated with the presence of a tyrosine at position M197. For 16 species, we found sequences for the M subunit and the C subunit. In 14 of these sequences, a tyrosine is present at position M197, suggesting such a correlation.

In the PSII RC, a redox potential of ~1100 mV is needed for the water splitting in the oxygen-evolving center (*67, 68*). It might well be that this redox potential is not achieved by the special pair, since it may not play the same role in charge separation in PSII as it does in the bRC. Experimental and theoretical studies suggested that the excitation is more distributed over the bound chlorophyll molecules in PSII

Table 6: Conservation of the Residues Suggested To Direct the Electron Transfer along the A Branch in the bRC of *B. viridis*[a]

		bRC		PSII RC			
function	residue	conservation (%)	exchanged to (%)	residue	conservation (%)	exchanged to (%)	*Rb. sphaeroides* numbering
favor	ArgL10	100.0		GluD1-10	80.1	S/Q/R/K/A/H/D/T (8.1/2.5/2.5/2.5/1.9/1.2/0.6/0.6)	ArgL10
$P^+Bph_A^-$	ArgL103	100.0		ArgD1-129	100.0		ArgL103
	ArgL109	98.0	K/N (1.0/1.0)	TyrD1-135	33.4	F/C (64.8/1.8)	ArgL109
	LysL110	100.0		ArgD1-136	96.4	P (3.6)	LysL110
	AspL155	96.0	D/V (3.0/1.0)	AsnD1-181	100.0		AspL155
	AsnL158	6.0	S (94.0)	IleD1-184	76.4	L/F (20.0/3.6)	SerL158
	AsnL166	34.0	H/Q (60.0/6.0)	AsnD1-191	100.0		AsnL166
	ArgL231	100.0		PheD1-273	100.0		ArgL231
	GluM171	93.8	Q/R/K/H (3.5/0.9/0.9/0.9)	PheD2-168	99.2	S/L (0.4/0.4)	GluM173
	AspM183	92.8	D/N/I/G (2.7/2.7/0.9/0.9)	ArgD2-180	100.0		AspM184
	AsnM193	98.2	S (1.8)	AsnD2-190	100.0		AsnM195
	ArgM226	98.2	L/G (0.9/0.9)	PheD2-223	96.2	Y (3.8)	ArgM228
	ArgM239	97.2	I/H (1.9/0.9)	GluD2-242	100.0		ArgM241
	ArgM245	98.0	S/P (1.0/1.0)	ThrD2-248	99.6	K (0.4)	ArgM247
	ArgM251	100.0		SerD2-254	100.0		ArgM253
	ArgM265	97.9	K (2.1)	PheD2-275	99.6	Y (0.4)	ArgM267
	AspM290	88.9	D (11.1)	GluD2-312	100.0		AspM292
oppose	GluL6	100.0		LeuD1-5	88.1	I/V/A/Q (8.8/1.3/1.3/0.6)	GluL6
$P^+Bph_A^-$	AspL23	98.4	D (1.6)	ThrD1-22	99.4	V (0.6)	AspL23
	GluL106	100.0		GluD1-132	98.2	D/G/V (0.6/0.6/0.6)	GluL106
	ArgL135	100.0		TyrD1-161	100.0		ArgL135
	HisL168	95.0	E (5.0)	LeuD1-193	99.4	Y (0.6)	HisL168
	AspM2	5.3	E/T/A/Q/S/Y/R/N (58.5/17.0/5.3/5.3/3.2/3.2/1.1/1.1)		–		GluM2
	ArgM190	57.0	T/I/K/V/N/L/A/D (13.4/10.7/7.2/5.4/2.7/1.8/0.9/0.9)	PheD2-188	99.6	L (0.4)	ValM192
	AspM230	44.0	E (55.0)	GluD2-227	84.3	D7G/N/Q/V (11.9/2.1/0.8/0.4/0.4)	GluM232
	GluM232	100.0		AlaD2-229	89.9	S/Y/P (5.5/3.3/1.3)	GluM234
	GluM244	99.0	D (1.0)	ValD2-247	99.6	I (0.4)	GluM246
	GluM261	97.9	R (2.1)	ArgD2-265	100.0		GluM263

[a] The amino acids are grouped into positive (R, K, and H), negative (D and E), polar (N, C, Q, S, T, and Y), and unpolar (A, I, L, P, and G) classes. The numbering refers to the bRC of *B. viridis*. In addition also, the numbering of the bRC of *Rb. sphaeroides* is given.

Table 7: Conservation of Functionally Important Residues of the Q_A Site and the Q_B Site[a]

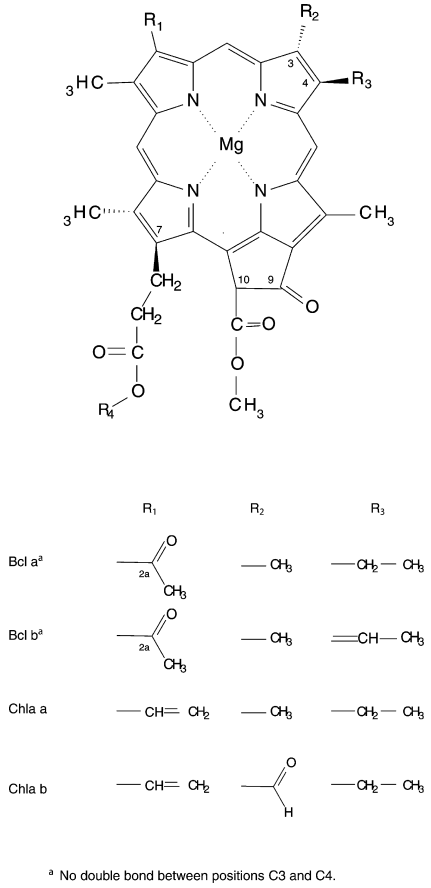
	bRC			PSII RC		
location	residue	conservation (%)	exchanged to (%)	residue	conservation (%)	exchanged to (%)
$Q_A$site	*HisM219*	100.0		HisD2-214	99.6	Y (0.4)
	AlaM248	88.4	G/S/C (7.4/3.2/1.1)	AlaD2-249	100.0	
	AlaM249	78.5	G/M/Q (15.1/5.4/1.1)	AsnD2-250	100.0	
	TrpM252	100.0		TrpD2-253	99.2	R (0.8)
	AlaM260	100.0		SerD2-262	98.3	A/P (1.3/0.4)
	IleM265	89.4	V/F (8.5/2.1)	LeuD2-267	99.6	S (0.4)
	HisM266	100.0		HisD2-268	99.6	Y (0.4)
$Q_B$site	*HisL190*	100.0		HisD1-215	100.0	
	GluL212	100.0		AlaD1-251	98.2	V (1.8)
	AspL213	60.0	N (40.0)	HisD1-252	99.4	Q (0.6)
	PheL216	95.8	Y/T (3.2/1.0)	PheD1-255	100.0	
	SerL223	100.0		SerD1-264	97.6	R/G (1.8/0.6)
	IleL224	64.5	V/N/S (33.3/1.1/1.1)	PheD1-265	100.0	
	IleL229	100.0		LeuD1-270	100.0	

[a] Marker residues are shown in italics. Residues for which only the backbone is involved in binding are shown in bold. The residue numbering of bRC and PSII RC refers to *Rb. sphaeroides* and *T. elongatus*, respectively.

(*16*). Recently, the accessory chlorophyll pigment of the A branch ($Cla_{D1\text{-}2}$; see Figure 1) was thought to be the true primary electron donor in PSII (*17*−*20*). Since bRC and PSII might have a different charge separation mechanism, the hydrogen bond pattern does not need to be conserved.

(ii) Hydrogen Bonding of the Other Pigment Molecules. Besides the special pair, each branch consists of an accessory chlorophyll and a pheophytin molecule (see Figure 1). Protein residues forming hydrogen bonds to the oxygen atoms of these pigments influence their binding (by hydrogen bonds to the C_7 phytyl and C_{10} carboxymethyl oxygen atoms) or tune their redox properties (via a hydrogen bond to the π-conjugated C_9 keto and C_{2a} acetyl oxygen atoms). The existence and conservation of hydrogen bonds to these pigment oxygen atoms are analyzed as described above. Figure 5 depicts all residues close to the oxygen atoms of the pigments. In Tables 4 and 5, the conservation of these residues is listed for the chlorophyll and pheophytin pigment molecules, respectively.

The first pigments that we consider are the accessory chlorophyll molecules. The C_7 phytyl oxygen atom of Bcl_{A2} in the A branch does not form a hydrogen bond to the bRC. The only residue which would be close enough is the conserved TyrM210, but the side chain of this residue is pointing away from the oxygen atoms. In the B branch of the bRC, among SerL178 and ThrL182, there is always one residue forming a hydrogen bond with the C_7 phytyl oxygen atoms of Bcl_{B2}. In contrast to the bRC, no hydrogen bond is



[a] No double bond between positions C3 and C4.

FIGURE 3: Chemical structure of (bacterio) chlorophyll. R_4 stands for the phytyl chain. (Bacterio) pheophytin has the same chemical structure but lacks the magnesium ion.

formed with the C_7 phytyl oxygen molecules of $Cla_{D1\text{-}2}$ or $Cla_{D2\text{-}2}$ of PSII. The C_{10} carboxymethyl oxygen atoms of Bcl_{A2} are singly hydrogen bonded to CysM203 in ~20% of the bRC sequences, whereas in the B branch of the bRC, this oxygen atom of Bcl_{B2} is hydrogen bonded with SerL178 in almost all sequences. In PSII, the C_{10} carboxymethyl oxygen atoms of $Cla_{D1\text{-}2}$ and $Cla_{D2\text{-}2}$ are singly hydrogen bonded (to SerD1-177) and not hydrogen bonded, respectively. The C_9 keto oxygen of Bcl_{A2}, which is part of the conjugated π-system, is hydrogen bonded to a water molecule close to GlyM203 in the bRC of *Rb. sphaeroides*. GlyM203 is not conserved (see Table 4), but it is exchanged with Met, Cys, Val, or Ala. In the bRC of *B. viridis*, in which a methionine is located at the position corresponding to M203, the water and thus the hydrogen bond are also observed. Thus, although GlyM203 is not conserved, the hydrogen bond to a water molecule seems to be conserved in all bRC proteins. In the B branch, HisL168 is close to the C_9 keto oxygen of Bcl_{B2}; however, the side chain of HisL168 is not pointing toward these oxygen atoms, and HisL168 is hydrogen bonded to the C_{2a} acetyl group of Bcl_{A1}. However, in analogy to the A branch, a water (close to MetL174) forms a hydrogen bond with the C_9 keto oxygen of Bcl_{B1} in the bRC of *Rb. sphaeroides*. The conserved residue MetL174 itself is more than 5 Å from the C_9 keto oxygen of Bcl_{B1}. A hydrogen bond between the C_9 keto oxygen of Bcl_{A1} and this water is likely to be conserved. In PSII, no hydrogen bonds to the C_9 keto oxygen atoms of the accessory chlorophyll are found. The C_{2a} acetyl oxygen atom of Bcl_{A2} forms a hydrogen bond with the conserved TyrL128 in bRC. In the B branch, a hydrogen bond is formed by TrpM157 with the C_2 keto oxygen to Bcl_{B2} in the bRC of *Rb. sphaeroides*, but this residue is not conserved.

The hydrogen bonding of the pheophytin molecules differs from that of the accessory chlorophyll molecules. The C_7 phytyl oxygen atoms are close to TrpL100 in the bRC. However, this residue forms a hydrogen bond to the C_{10} carboxymethyl oxygen atoms of Bph_A. Thus, no hydrogen bond is formed to the C_7 phytyl oxygen atoms of the Bph pigments in bRC. Also, in PSII, no hydrogen bonds are formed to the C_7 phytyl oxygen atoms of the pheophytin pigments. The C_{10} carboxymethyl oxygen atom of Bph_A is close to ThrM255, TrpL100, and GluL104 in the bRC. Although these residues are not strictly conserved, they are exchanged only with amino acids, which are able to form a hydrogen bond. However, the hydroxyl group of ThrM255 is oriented away from the oxygen and is likely to be hydrogen bonded with the backbone of M251. A hydrogen bond is formed between TrpL100 and the C_{10} methylester oxygen atom of Bph_A. This hydrogen bond was also detected spectroscopically (*69*−*71*). In the B branch of the bRC, the C_{10} carboxymethyl oxygen atoms of Bph_B forms a hydrogen bond with the conserved TrpM129. In PSII, it is likely that the C_{10} carboxymethyl oxygen atoms of the Phe pigments are bound as in bRC proteins. The C_9 keto oxygen of Bph_A forms a hydrogen bond with TrpL100 and GluL104 in the bRC of *Rb. sphaeroides*. Although GluL104 is not strictly conserved, it is only exchanged with Gln, which can also form a hydrogen bond. The hydrogen bond between GluL104 and the C_9 keto oxygen of Bph_A is also determined experimentally (*69*−*71*). In the B branch, the C_9 keto oxygen of Bph_B forms a hydrogen bond with ThrM133 in most bRC sequences. Moreover, a hydrogen bond donor is found in some bRC sequences at positions L219 and M146. Thus, the C_9 keto oxygen of Bph_B is most probably in all bRC proteins singly hydrogen bonded. In PSII, the C_9 keto oxygen atoms of $Pheo_{D2}$ and $Pheo_{D1}$ are hydrogen bonded by two residues (TyrD1-126 and GlnD1-130; see Table 5) and one residue (GlnD2-129), respectively, resembling the situation in the bRC. In addition, a second hydrogen bond to the C_9 keto oxygen of $Pheo_{D1}$ (by AsnD2-142) seems likely. The C_{2a} acetyl oxygen atom of Bph_A is close to the conserved residue TyrM210 in the bRC, but since this residue is pointing in the wrong direction, it is not likely that the C_{2a} acetyl oxygen of Bph_A is hydrogen bonded. In the B branch, the C_{2a} acetyl oxygen of Bph_B is not hydrogen bonded.

Our analysis shows a larger number of hydrogen bonds to the C_7 phytyl and C_{10} carboxymethyl oxygen atoms of the accessory bacteriochlorophyll and the bacteriopheophytin pigments of the B branch compared to the pigments of the A branch in bRC proteins, but in PSII, the number of hydrogen bonds is higher for the A branch than for the B branch. These hydrogen bonds influence most likely the binding and positioning of the pigments but less the redox potential. In the bRC, more hydrogen bonds to oxygen atoms that are part of the conjugated π-system (C_9 keto and C_{2a} acetyl oxygen) are formed in the A branch than in the B branch. Since these hydrogen bonds influence the redox potential of the pigments, they may be partially responsible for directing the electron transfer along the A branch. In PSII RC proteins, however, the number of hydrogen bonds to the π-conjugated oxygen atoms is the same in both branches,

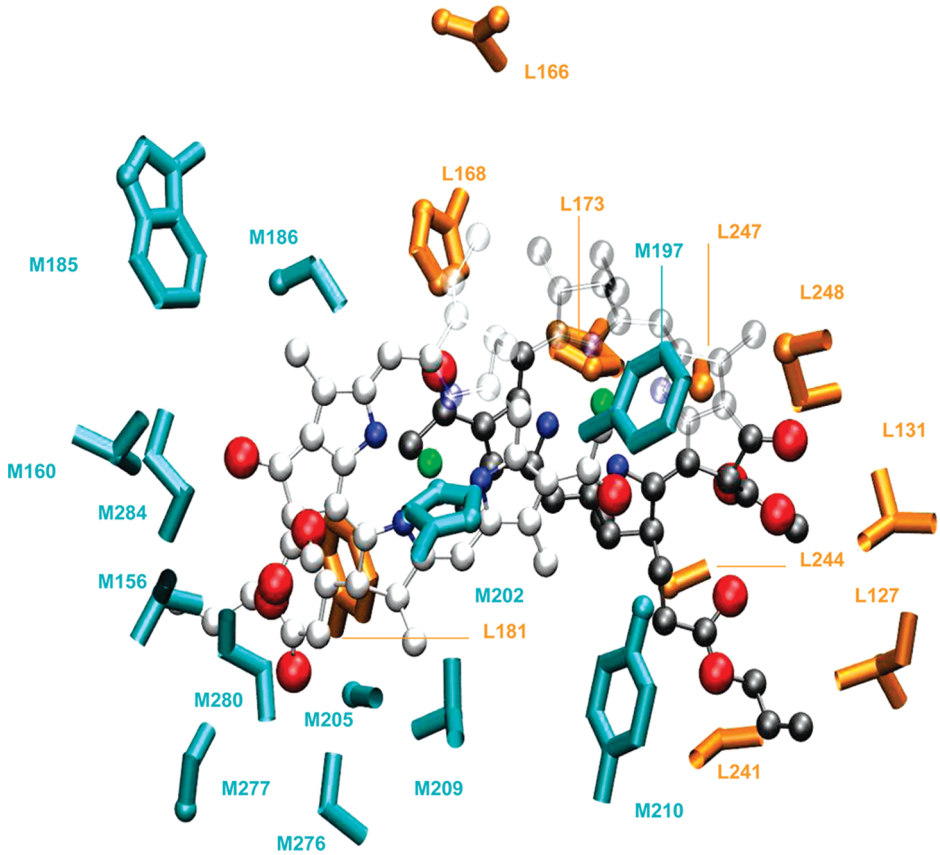


FIGURE 4: Special pair (shown as balls and sticks) surrounded by residues (shown as sticks). The residues are colored corresponding to their subunit (L, cyan; and M, orange) and are depicted using the bond representation. The non-carbon atoms of the amino acid residues are shown as spheres. This figure was generated with vmd (*80*) using PDB structure 2J8C (*49*) of the bRC from *Rb. sphaeroides*.

which may again reflect the different charge separation mechanism in PSII.

Directing the Electron Transfer along the A Branch. Continuum electrostatic calculations showed that the electrostatic potential of the pigments is different in the A branch and the B branch of the bRC from *B. viridis*, which might be a reason the electron transfer occurs along the A branch (*26*). In this earlier study by Gunner et al. (*26*), residues which favor (ArgL10, ArgL103, ArgL109, AspL155, AspL158, AspL166, ArgL231, GluM171, AspM183, AsnM193, ArgM226, ArgM239, ArgM245, ArgM251, ArgM265, and ArgM290) or oppose (GluL6, AspL23, GluL106, ArgL135, HisL168, AspM2, ArgM190, AspM230, GluM232, GluM244, and GluM261) the formation of the $P^+Bph_A^-$ state were identified. To the best of our knowledge, no continuum electrostatic calculations identifying residues which favor or oppose the formation of the $P^+Bcl_{A2}^-$ state, which proceeds the $P^+Bph_A^-$ state, have been reported. However, the formation of the $P^+Bph_A^-$ state is considered to be the key event for directing the electron transfer along the A branch. In this study, we analyze the degree of conservation of the residues that were proposed by Gunner et al. (*26*) to be involved in directing the electron transfer along the A branch by favoring the $P^+Bcl_{A2}^-$ state. A residue is considered functionally conserved when it is conserved or exchanged with an amino acid with a similar character (negative, positive, polar, and unpolar); i.e., the residue keeps its charge or polarity. The identified residues of the L and M subunits are depicted in Figure 6. The conservation of these residues is listed in Table 6. In the following, the amino acid character of the residues and numbering refer to those of the bRC of *B. viridis*, since the electrostatic calculations were performed on this bRC. Except AsnL166, GluM171, and AspM183, all residues favoring the formation of $P^+Bph_A^-$ are conserved or are exchanged with functionally similar residues in the bRC. GluM171 and AspM183 are only exchanged in few sequences, and an exchange with an unpolar residue is never observed. AsnL166 was discussed already above, because of its influence on the hydrogen bond between HisL168 and the special pair in the bRC (*57*). AsnL166 is exchanged with a histidine in 60% of the analyzed L subunit sequences. A histidine at position L166 can also stabilize this hydrogen bond. Thus, the exchange of AsnL166 with His is not likely to alter the stabilization of $P^+Bph_A^-$ significantly. The residues destabilizing the $P^+Bph_A^-$ state are not conserved to the same extent in the bRC (see Table 6). For example, AspM2 and ArgM190 are exchanged with amino acids with different character (positive, negative, polar, and unpolar). The observed degree of conservation of residues favoring the $P^+Bcl_{A2}^-$ state could in addition arise from the fact that these residues might be involved in the interaction with other proteins such as cytochrome or are important for protein folding. However, several of these residues are functionally conserved (GluL6, AspL23, GluL106, ArgL135, AspM230, GluM232, and GluM244; see Table 6). Most of the residues proposed by Gunner et al. (*26*) to cause a different electric field in the two branches are functionally conserved in bRC proteins. It is therefore likely that these residues strongly favor the electron transfer along the A branch in all bRC proteins.

In contrast, in PSII RC proteins, many of these residues are exchanged with an amino acid with a different character (see Table 6). Thus, either the electrostatic potential is not different along the two branches in PSII, or the differences are caused by other residues. Moreover, the charge separation in the PSII

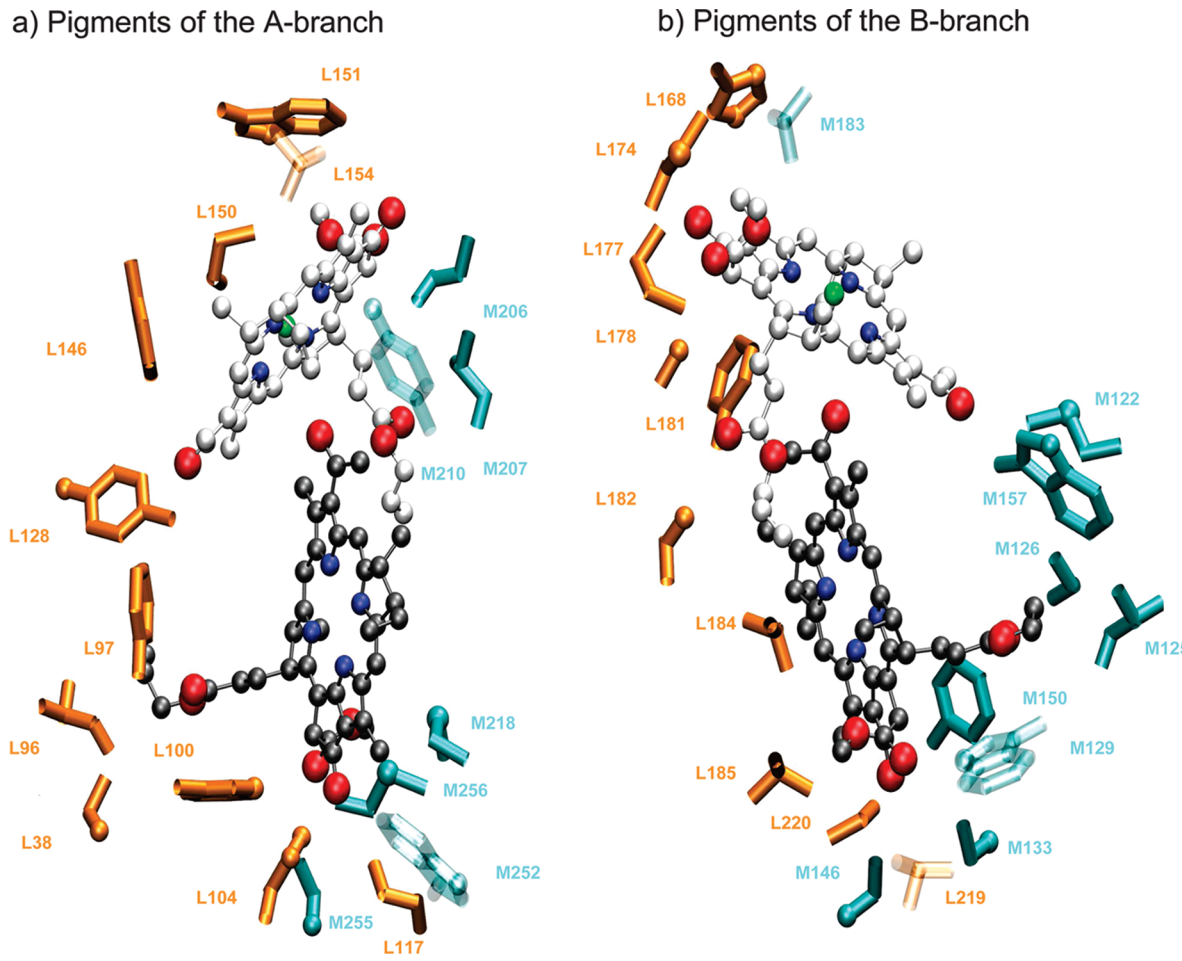


FIGURE 5: Accessory bacteriochlorophyll (white) and bacteriopheopytin (gray) molecules (shown as balls and sticks) are shown for the A branch (a) and the B branch (b). The residues that surround the pigments (shown as sticks) are colored corresponding to their subunit (L, cyan; and M, orange). The residues are colored corresponding to their subunit (L, cyan; and M, orange) and are depicted using the bond representation. The non-carbon atoms of the amino acid residues are shown as spheres. This figure was generated with vmd (*80*) using PDB structure 2J8C (*49*) of the bRC from *Rb. sphaeroides*.

RC may not occur at the special pair, but at the accessory chlorophyll of the A branch (Cla_{D1-2}) (*16*−*20*). Thus, it is very likely that in PSII the tuning of the redox properties and direction of the charge transfer along the A branch is not determined by the same mechanism as in the bRC.

Q_A and Q_B Binding Sites. In the bRC, the different character of the Q_A and Q_B binding sites is mainly determined by the amino acids at the symmetric positions M248/M249 (Q_A site) and L212/L213 (Q_B site). L212 and L213 are charged or polar residues, whereas M248 and M249 are unpolar. Several residues are involved in binding and stabilizing Q_A (HisM219, TrpM252, and AlaM260) and Q_B (HisL190, PheL216, SerL223, Ile224, and IleM229). Moreover, it has been shown that IleM265 and HisM266 are important for adjusting the midpoint potential of Q_A (*72, 73*). We analyzed the conservation of the different character of the two binding sites in bRC and PSII proteins. The two quinone binding sites of the bRC and of the PSII RC are depicted in Figure 7. The results of this conservation analysis are listed in Table 7. Most of the listed residues are conserved or are only exchanged with similar residues in the bRC. Residues HisM219, TrpM252, HisL190, PheL216, and SerL223, which are directly involved in quinone binding, are conserved in type II RC proteins. The conservation of IleL224 and IleL229 is not critical for Q_B binding, since only the backbone of these residues forms a hydrogen bond with Q_B. It has been shown that the mutation of AlaM260 to a bulky amino acid leads to weak binding of Q_A in the bRC of *Rb. sphaeroides* (*74*). AlaM260 is always a small amino acid in the bRC as well as in PSII. IleM265, which influences the redox potential of Q_A in the bRC (*72*), is always a Leu in PSII. HisM266 also influences the redox potential of Q_A and is conserved in all type II RC proteins. This histidine is also involved in the coordination of the non-heme iron.

In the bRC, GluL212 and AspL213 are known to be important for transfer of the proton to Q_B (*75, 76*). It is known that at position L213 either Asp (like in the bRC of *Rb. sphaeroides*) or Asn (like in the bRC of *B. viridis*) is found (*77, 78*). These two alternative amino acids are the only amino acids observed at position L213 in our PA of the bRC. GluL212 is strictly conserved in the bRC. Residues AlaM248 and AlaM249 are not strictly conserved in the analyzed bRC sequences; however, they are exchanged with amino acids with similar character. Thus, the different character of the Q_A site and the Q_B site is conserved in the bRC. In PSII, an Ala and a His, both conserved, are found at positions L212 and L213, respectively, making the Q_B site less charged than in the bRC. Therefore, it is likely that the transfer of the proton to Q_B is different in PSII RC proteins and bRC proteins (*79*). At positions M248 and M249, an Ala (D2-249) and an Asn (D2-250) are found, respectively, which makes the Q_A site more polar than in

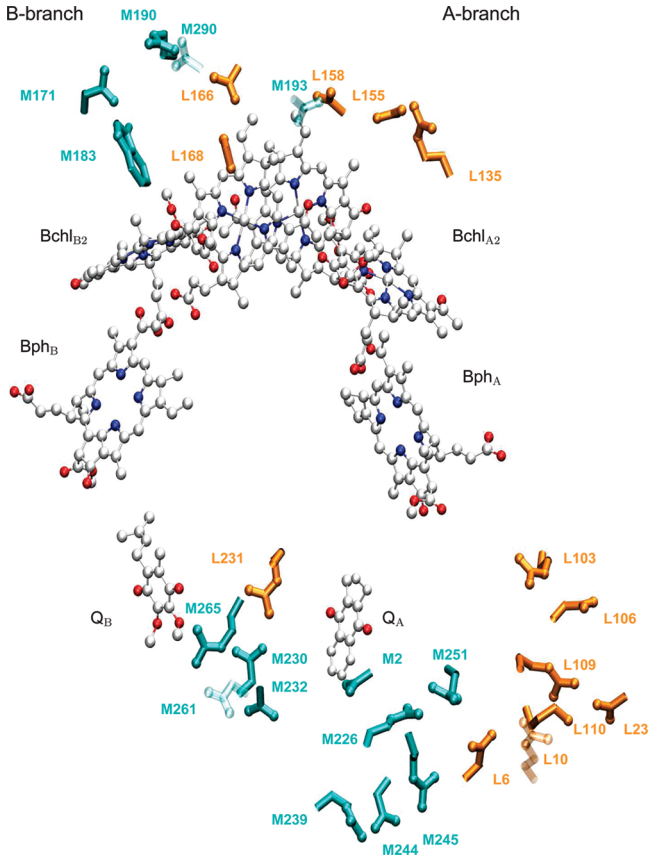


FIGURE 6: Residues that direct the electron transfer along the A branch. The pigments and the quinones are shown in a ball-and-stick mode. Protein residues are shown as sticks. The residues are colored corresponding to their subunit (L, cyan; and M, orange) and are depicted using the bond representation. The non-carbon atoms of the amino acid residues are shown as spheres. This figure was generated with vmd (*80*) using PDB structure 2I5N (*81*) of the bRC of *B. viridis*.

the bRC. The examination of the crystal structure of PSII shows that the polar side chain of AsnD2-250 points away from Q_A (see Figure 7b), and thus, the influence of the polar side chain on the properties of Q_A might be weak. However, the slightly more polar environment of Q_A in PSII proteins compared to the situation in the bRC could explain why a doubly reduced and protonated Q_A site occurs during photoinhibition in PSII (*21*).

On the basis of our analysis, we are able to show that residues directly involved in the binding of Q_A and Q_B are conserved in all type II RC proteins, but the character of the two quinone binding sites differs slightly between PSII RC and bRC proteins, leading to the possibility that Q_A is protonated in PSII under stress while a protonated Q_A was never observed in bRC.

CONCLUSIONS

The bacterial and PSII RCs perform the same biological task and thus share many similarities. Nevertheless, these proteins differ in their detailed molecular mechanism. In this study, we analyzed the similarities and dissimilarities of the two proteins at the level of their sequences. Due to the low level of sequence identity of the analyzed PSII and bRC proteins, profile hidden Markov models are needed to align their sequences properly. A progressive alignment failed to align functionally important residues. The profile hidden

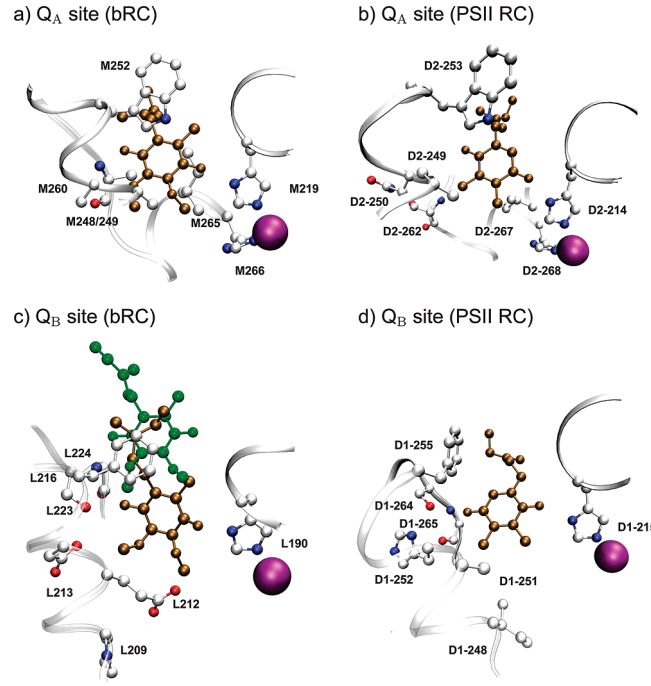


FIGURE 7: Q_A and Q_B sites with important residues. In all figures, important residues for binding as well as the non-heme iron (purple) and the quinone (brown) are shown. (a) Q_A binding site of the bRC of *Rb. sphaeroides*. (b) Q_A binding site of the PSII RC of *T. elongatus*. (c) Q_B site of the bRC of *Rb. sphaeroides* shown with the quinone in its proximal (brown) and distal (green) positions. (d) Q_B site of the PSII RC of *T. elongatus*. This figure was generated with vmd (*80*) using PDB structures 2J8C (*49*) and 2AXT (*15*).

Markov model was constructed with the help of the available structural information. This is to the best of our knowledge the first time that a profile hidden Markov model is constructed for each subunit (one for the L/D1 subunit and one for the M/D2 subunit). We show that residues involved in binding of Q_A and Q_B are conserved, but residues forming hydrogen bonds to the pigments are not. Residues involved in midpoint potential tuning of the special pair and in determining the redox properties of the A branch and the B branch are conserved in the bRC but have a different amino acid character in PSII. The strategy followed in this paper to generate a multiple-sequence alignment using a profile hidden Markov model constructed with the help of structural information can be seen as an illustrative example of comparing proteins with a very low level of sequence identity.

SUPPORTING INFORMATION AVAILABLE

A figure showing examples of misaligned sequences, a table containing all sequences for the construction of the pHMM for the L/D1 subunit and the M/D2 subunit, a table containing all strictly conserved residues in the L, M, D1, and D2 subunits, a table containing all strictly conserved residues in the RC, all sequences used for the PA, and a zip file containing the two pHMM files, the MSA file of the two misalignments, and the PA alignment files for the validation of the pHMMs and for the conservation analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Heathcote, P., Fyfe, P. K., and Jones, M. R. (2002) Reaction centres: The structure and evolution of biological solar power. *Trends Biochem. Sci. 27*, 79–87.
2. Rutherford, A. W., and Faller, P. (2001) The heart of photosynthesis in glorious 3D. *Trends Biochem. Sci. 26*, 341–344.
3. Cogdell, R. J., and Lindsay, J. G. (2000) The structure of photosynthetic complexes in bactera and plants: An illustration of the importance of protein structure to the future development of plant science. *New. Phytol. 145*, 167–196.
4. Baymann, F., Brugna, M., Mühlenhoff, U., and Nitschke, M. (2001) Daddy, where did (PS)I come from. *Biochim. Biophys. Acta 1507*, 291–301.
5. Meyer, T. E. (1994) Evolution of photosynthetic reaction centers and light harvesting chlorophyll proteins. *BioSystems 33*, 167–175.
6. Blankenship, R. E. (1992) Origin and early evolution of photosynthesis. *Photosynth. Res. 33*, 91–111.
7. Margulies, M. M. (1991) Sequences similarity between Photosystem I and II. Identification of a Photosystem I reaction center transmembrane helix that is similar to transmembrane helix IV of the D2 subunit of Photosystem II and the M subunit of the non-sufur purple and flexible green bacteria. *Photosynth. Res. 29*, 133–147.
8. Allen, J. P., and Williams, J. C. (1998) Photosynthetic reaction centers. *FEBS Lett. 438*, 5–9.
9. Svensson, B., Etchebest, C., Tuffery, P., vanKan, P., Smith, J., and Styring, S. (1996) A model for the photosystem II reaction center core including the structure of the primary donor P_{680}. *Biochemistry 35*, 14486–14502.
10. Olson, J. M., and Blankenship, R. E. (2004) Thinking about the evolution of photosynthesis. *Photosynth. Res. 80*, 373–386.
11. Mulkidjanian, A. Z., Koonin, E. V., Makarova, K. I., Mekhedov, S. L., Sorokin, A., Wolf, Y. I., Dufresne, A., Partensky, F., Burd, H., Kaznadzey, D., Haselkorn, R., and Galperin, M. Y. (2006) The cyanobacterial genome core and the origin of photosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 103*, 13126–13131.
12. Michel, H., and Deisenhofer, J. (1988) Relevance of the photosynthetic reaction center from purple bacteria to the structure of photosystem II. *Biochemistry 27*, 1–6.
13. Rost, B. (1997) Protein structures sustain evolutionary drift. *Folding Des. 2*, S19−S24.
14. Blankenship, R. E. (1994) Protein structure, electron transfer and evolution of prokaryotic photosynthetic reaction centers. *Antonie van Leeuwenhoek 65*, 311–329.
15. Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005) Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. *Nature 438*, 1040–1044.
16. Durrant, J. R., Klug, D. R., Kwa, S. L., vanGrondelle, R., Porter, G., and Dekker, J. P. (1995) A multimer model for P680, the primary electron donor of photosystem II. *Proc. Natl. Acad. Sci. U.S.A. 92*, 4798–4802.
17. Groot, M. L., Pawlowicy, N. P., vanWilderen, L. J. G. W., Breton, J., vanStokkum, I. H. M., and vanGrondelle, R. (2005) Initial electron donor and acceptor in isolated photosystem II reaction centers identified with femtosecond mid-IR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A. 102*, 13087–13092.
18. Holzwarth, A. R., Müller, M. G., Reus, M., Novwacyzk, M., Sander, J., and Rögner, M. (2006) Kinetics and mechanism of electron transfer in intact phosystem II and in the isolated reaction center: Pheophytin is the primary electron acceptor. *Proc. Natl. Acad. Sci. U.S.A. 103*, 6895–6900.
19. Raszewski, G., Saenger, W., and Renger, T. (2005) Theory of Optical Spectra of Photosystem II Reaction Centers: Location of the Triplet State and the Idenity of the Primary Electron Donor. *Biophys. J. 88*, 986–998.
20. Renger, T., and Schlodder, E. (2008) The Primary Electron Donor of Photosystem II of the Cyanobacterium *Acaryochloris marina* Is a Chlorophyll *d* and the Water Oxidation Is Driven by a Chlorophyll *a*/Chlorophyll *d* Heterodimer. *J. Phys. Chem. B 112*, 7351–7354.
21. Krieger-Liszkay, A. (2004) Singlet oxygen production in photosynthesis. *J. Exp. Biol. 56*, 337–346.
22. Noguchi, T. (2002) Dual role of triplet localization on the accessory chlorophyll in the photosystem II reaction center: Photoprotection and photodamage of the D1 protein. *Plant Cell Physiol. 43*, 1112–1116.
23. Laible, P. D., Morris, Z. S., Thurnauer, M. C., Schiffer, M., and Hanson, D. K. (2003) Inter- and intraspecific variation in excited-state triplet engery transfer rates in reaction centers of photosynthetic bacteria. *Photochem. Photobiol. 78*, 114–123.
24. McAuley, K. E., Fyfe, P. K., Codgell, R. J., Isaacs, N. W., and Jones, M. R. (2000) X-ray crystal structure of the YM210W mutant reaction centre from *Rhodobacter sphaeroides*. *FEBS Lett. 467*, 285–290.
25. vanBrederode, M. E., Ridge, J. P., vanStokkum, I. H. M., vanMourik, F., Jones, M. R., and vanGrondelle, R. (1998) On the efficiency of energy transfer and the different pathways of electron transfer in mutant reaction centers of *Rhodobacter sphaeroides*. *Photosynth. Res. 55*, 141–146.
26. Gunner, M. R., Nicholls, A., and Honig, B. (1996) Electrostatic potentials in *Rhodopseudomonas viridis* reaction centers: Implications for the driving force and directionality of electron transfer. *J. Phys. Chem. 100*, 4277–4291.
27. Kirmaier, C., Laible, P. D., Hanson, D. K., and Holten, D. (2003) B-side charge separation in bacterial photosynthetic reaction center: Nanosecond time scale electron transfer from H_B^- to Q_B. *Biochemistry 42*, 2016–2024.
28. Chang, C., Ekins, S., Bahadduri, P., and Swaan, P. W. (2006) Pharmacophore-based discovery of ligands for drug transporters. *Adv. Drug Delivery Rev. 58*, 1431–1450.
29. Okamura, M. Y., and Feher, G. (1992) Proton transfer in reaction centers from photosynthetic bacteria. *Annu. Rev. Biochem. 61*, 861–896.
30. Brudler, R., deGroot, H. J. M., vanLiemt, W. B. S., Gast, P., Hoff, A. J., Lugtenburg, J., and Gewert, K. (1995) FTIR spectroscopy shows weak symmetric hydrogen bonding of the Q_B carbonyl groups in *Rhodobacter sphaeroides* R26 reaction centres. *FEBS Lett. 370*, 88–92.
31. Takahashi, E., Wells, T. A., and Wraight, C. A. (2001) Protein control of the redox potential of the primary quinone acceptor in reaction centers from *Rhodobacter sphaeroides*. *Biochemistry 40*, 1020–1028.
32. Fritzsch, G., Koepke, J., Diem, R., Kuglstatter, A., and Baciou, L. (2002) Charge separation induces conformational changes in the photosynthetic reaction centre of purple bacteria. *Acta Crystallogr. 58*, 1660–1663.
33. Thompson, J. D., Higgins, D. G., and Gibson, J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673–4680.
34. Rost, B. (1999) Twilight zone of protein sequence alignments. *Protein Eng. 12*, 85–94.
35. Sauder, J. M., Arthur, J. W., and Dunbrack, R. L., Jr. (2000) Large-scale comparison of protein sequence alignment algorithms with structure alignments. *Proteins 40*, 6–22.
36. Shatsky, M., Dror, O., Schneidman-Duhovny, D., Nussinov, R., and Wolfson, H. J. (2004) BioInfo3D: A suite of tools for structural bioinformatics. *Nucleic Acids Res. 32*, W503−W507.
37. Pei, J., Kim, B. H., and Grishin, N. V. (2008) PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Res. 36*, 2295–2300.
38. Wallance, I. M., Blackshields, G., and Higgins, D. G. (2005) Multiple sequence alignments. *Curr. Opin. Struct. Biol. 15*, 261–266.
39. O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D. G., and Notredame, C. (2004) 3DCoffee: Combining protein sequences and structures within multiple sequence alignments. *J. Mol. Biol. 340*, 385–395.
40. Shatsky, M., Nussinov, R., and Wolfson, H. J. (2006) Optimization of multiple-sequence alignment baed on muliple structure alignment. *Proteins 62*, 209–217.
41. Dubin, R., Eddy, S., Krogh, A., and Mitchison, G. (1998) *Biological sequence analysis*, Cambridge University Press, New York.
42. Eddy, S. R. (1996) Hidden Markov Models. *Curr. Opin. Struct. Biol. 6*, 361–365.
43. Eddy, S. R. (1998) Profile hidden Markov models. *Bioinf. Rev. 14*, 755–763.
44. Birney, E. (2001) Hidden Markov models in biological sequence analysis. *J. Res. Dev. 34*, 449–454.
45. Mount, D. W. (2001) *Bioinformatics. Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, Plainview, NY.
46. Durbin, R., Eddy, S., Krogh, A., and Mitchison, G. (1998) *Biological sequence analysis: Probabilistic models of proteins and nucleic acids*, Cambridge University Press, New York.
47. Bernardes, J. S., Dävila, A. M. R., Costa, V. S., and Zaverucha, G. (2007) Improving model construction of profile HMMs for

remote homology detection through structural alignment. *BMC Bioinf. 8*, 435–447.
48. Wolfson, H., Shatsky, M., Schneidman-Duhovny, D., Dror, O., Schlman-Peleg, A., Ma, B., and Nussinov, R. (2005) From structure to function: Methods and applications. *Curr. Protein Pept. Sci. 2005*, 171–183.
49. Koepke, J., Krammer, E. M., Klingen, A. R., Sebban, P., Ullmann, G. M., and Fritzsch, G. (2007) PH Modulates the Quinone Position in the Photosynthetic Reaction Center from *Rhodobacter sphaeroides* in the Neutral and Charge Separated States. *J. Mol. Biol. 371*, 396–409.
50. Deisenhofer, J., Epp, O., Sinning, I., and Michel, H. (1995) Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodopseudomonas viridis*. *J. Mol. Biol. 246*, 429–457.
51. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol. 215*, 403–410.
52. Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Geer, L. Y., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., Maglott, D. R., Ostell, J., Miller, V., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R. L., Tatusova, T. A., Wagner, L., and Yaschenko, E. (2006) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res. 34*, D5−D12.
53. Sonnhammer, E. L. L., Eddy, S. R., and Durbin, R. (1997) Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins 28*, 405–420.
54. Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marschall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res. 34*, D247−D251.
55. Kuglstatter, A., Hellwig, P., Fritzsch, G., Wachtveitl, J., Oesterhelt, D., Mäntele, W., and Michel, H. (1999) Identification of a hydrogen bond in the Phe M197 → Tyr mutant reaction center of the photosynthetic purple bacterium *Rhodobacter sphaeroides* by X-ray crystallography and FTIR spectroscopy. *FEBS Lett. 463*, 169–174.
56. Lin, X., Murchison, H. A., Nagarajan, V., Parson, W. W., Allen, J. P., and Williams, J. C. (1994) Specific alteration of the oxidation potential of the electron donor in reaction centers from *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci. U.S.A. 91*, 10265–10269.
57. Ivancich, A., and Mattioli, T. A. (1997) Influence of Asn/His L166 on the hydrogen-bonding pattern and redox potential of the primary donor of purple bacterial reaction center. *Biochemistry 36*, 3027–3036.
58. Allen, J. P., Artz, K., Lin, X., Williams, J. C., Ivancich, A., Albouny, D., Mattioli, T. A., Fetsch, A., Kuhn, M., and Lubitz, W. (1996) Effects of hydrogen bonding to a bacteriochlorophyll-bacteriopheophytin dimer in reaction centers from *Rhodobacter sphaeroides*. *Biochemistry 35*, 6612–6619.
59. Ridge, J. P., Fyfe, P. K., McAuley, K. E., vanBrederode, M. E., Robert, B., vanGrondelle, R., Isaacs, N. W., Cogdell, R. J., and Jones, M. R. (2000) An examination of how structural changes can affect the rate of electron transfer in mutated bacterial photoreaction centre. *Biochem. J. 351*, 467–478.
60. Muegge, I., Apostolakis, J., Ermler, U., Fritzsch, G., Lubitz, W., and Knapp, E. W. (1996) Shift of the special pair redox potential: Electrostatic energy computations of mutants of the reaction center from *Rhodobacter sphaeroides*. *Biochemistry 35*, 8359–8370.
61. Reimers, J. R., Hughes, J. M., and Hush, N. S. (2000) Modeling the bacterial photosynthetic reaction center 3: Interpretation of effects of site-directed mutagenesis on the special-pair midpoint potential. *Biochemistry 39*, 16185–16189.
62. Mattioli, T. A., Gray, K. A., Luty, M., Oesterhelt, D., and Robert, B. (1991) Resonance Raman characterization of *Rhodobacter sphaeroides* reaction centers bearing site-directed mutations at tyrosine M210. *Biochemistry 30*, 1715–1722.
63. Ivancich, A., Kobayashi, M., Drepper, F., Fathir, I., Saito, T., Nozawa, T., and Mattioli, T. A. (1996) Hydrogen-bond interactions of the primary donor of the photosynthetic purple sulfur bacterium *Chromatium tepidum*. *Biochemistry 35*, 10529–10538.
64. Wachveitl, J., Farchaus, J. W., Das, R., Lutz, M., Robert, B., and Mattioli, T. A. (1993) Structure, spectroscopic, and redox properties of *Rhodobacter sphaeroides* reaction centers bearing point mutations near the primary electron donor. *Biochemistry 32*, 12875–12886.
65. Dracheva, S. M., Drachev, L. A., Konstantinov, A. A., Semenov, A. Y., Skulachev, V. P., Arutjunjan, A. M., Shuvalov, V. A., and Zaberezhaya, S. M. (1988) Electrogenic steps in the redox reactions catalyzed by photosynthetic reaction-centre complex. *Eur. J. Biochem. 171*, 253–264.
66. Ullmann, G. M., Kloppmann, E., Essigke, T., Krammer, E. M., Klingen, A. R., Becker, T., and Bombarda, E. (2008) Investigating the mechansims of photosynthetic proteins using continuum electrostatics. *Photosynth. Res. 97*, 33–53.
67. Ishikita, H., Loll, B., Biesiadka, J., Saenger, W., and Knapp, E. W. (2005) Redox potentials of chlorophylls in the photosystem II reaction center. *Biochemistry 44*, 4118–4124.
68. vanGorkom, H. J. (1985) Electron transfer in photosystem II. *Photosynth. Res. 6*, 97–112.
69. Mäntele, W. G., Wollenweber, A. M., Nabedryk, E., and Breton, J. (1988) Infrared spectroelectrochemistry of bacteriochlorophylls and bacteriopheophytins: Implication for the binding of the pigments in the reaction center from phtosythetic bacteria. *Proc. Natl. Acad. Sci. U.S.A. 85*, 8468–8472.
70. Yeates, T. O., Komiya, H., Chorine, A., Rees, D. C., Allen, J. P., and Feher, G. (1988) Structure of the reaction center from *Rhodobacter sphaeroides* R-26 and 2.4.1: Protein-cofactor (bacteriochlorophyll, bacteriopheophytin, and carotenoid) interaction. *Proc. Natl. Acad. Sci. U.S.A. 85*, 7993–7997.
71. Ermler, U., Fritzsch, G., Buchanan, S. K., and Michel, H. (1994) Structure of the photosynthetic reaction centre from *Rhodobacter sphaeroides* at 2.65 Å resuoltion: Cofactors and protein-cofactor interactions. *Structure 2*, 925–936.
72. Wells, T. A., Takahashi, E., and Wraight, C. A. (2003) Primary quinone (Q_A) binding site of bacterial photosynthetic reaction centers: Mutations at resude M265 probed by FTIR spectroscopy. *Biochemistry 42*, 4064–4074.
73. Cheap, H., Tandori, J., Derrien, V., Benoit, M., deOliveira, P., Koepke, J., Lavergne, J., Maróti, P., and Sebban, P. (2007) Evidence for delocalized anticooperative flash induced proton binding as revealed by mutatns at the M266His iron ligand in bacterial reaction centers. *Biochemistry 46*, 4510–4521.
74. Ridge, J. P., vanBrederode, M. E., Goodwin, M. G., vanGrondelle, R., and Jones, M. R. (1999) Mutations that modify or exclude binding of Q_A ubiquinone and carotenoid in the reaction center from *Rhodobacter sphaeriodes*. *Photosynth. Res. 59*, 9–26.
75. Paddock, M. L., Rongley, S. H., Feher, G., and Okamura, M. Y. (1989) Pathway of proton transfer in bacterial reaction centers: Replacement of glutamic acid 212 in the L subunit by glutamine inhibits quinone (secondary acceptor) turnover. *Proc. Natl. Acad. Sci. U.S.A. 86*, 6602–6606.
76. Takahashi, E., and Wraight, C. A. (1990) A crucial role for Asp L213 in the proton transfer pathway to the secondary quinone of reaction centers from *Rhodobacter sphaeroides*. *Biochim. Biophys. Acta 1020*, 107–111.
77. Rongley, S. H., Paddock, M. L., Feher, G., and Okamura, M. Y. (1993) Pathway of proton transfer in bacterial reaction centers: Second-site mutation Asn-M44 → Asp restores electron and proton transfer in reaction centers from the photosynthetically deficient Asp-L213 → Asn mutant of *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci. U.S.A. 90*, 1325–1329.
78. Paddock, M. L., Senft, M. E., Graige, M. S., Rongey, S. H., Turanchik, T., Feher, G., and Okamura, M. Y. (1998) Characterization of second site mutations show that fast proton transfer to Q_B^- is restored in bacterial reaction centers of *Rhodobacter sphaeroides* containing the Asp-L213 → Asn lesion. *Photosynth. Res. 58*, 281–291.
79. Ishikita, H., and Knapp, E. W. (2005) Control of Quinone Redox Potentials in Photosystem II: Electron Transfer and Photoprotection. *J. Am. Chem. Soc. 127*, 14714–14720.
80. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33–38.
81. Li, L., Mustafi, D., Fu, Q., Tereshko, V., Chen, D. L., Tice, J. D., and Ismagilov, R. F. (2006) Nanoliter microfluidic hydrid method for simultaneous screening and optimzation validated with crystallization of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A. 103*, 19243–19248.

BI802033K