Gunner, M. R., Tiede, D. M., Prince, R. C., & Dutton, P. L. (1982) in *Function of Quinone in Energy-Conserving Systems* (Trumpower, B. L., Ed.) pp 265–269, Academic Press, New York.
Gunner, M. R., Braun, B. S., Bruce, J. M., & Dutton, P. L. (1985) in *Antennas and Reaction Centers of Photosynthetic Bacteria* (Michel-Beyerle, M. E., Ed.) pp 298–305, Springer, Berlin.
Ikegami, I., & Katoh, S. (1989) *Plant Cell Physiol. 30*, 175–182.
Ikegami, I., Sêtif, P., & Mathis, P. (1987) *Biochim. Biophys. Acta 894*, 414–422.
Itoh, S., & Iwaki, M. (1988) *Biochim. Biophys. Acta 934*, 32–38.
Itoh, S., & Iwaki, M. (1989a) *FEBS Lett. 243*, 47–52.
Itoh, S., & Iwaki, M. (1989b) *FEBS Lett. 250*, 441–447.
Itoh, S., & Iwaki, M. (1991) *Biochemistry* (preceding paper in this issue).
Itoh, S., Iwaki, M., & Ikegami, I. (1987) *Biochim. Biophys. Acta 893*, 508–516.
Iwaki, M., & Itoh, S. (1989) *FEBS Lett. 256*, 11–16.
Iwaki, M., & Itoh, S. (1991) in *Advances in Chemistry Series No. 228, Electron Transfer in Inorganic, Organic, and Biological Systems* (Bolton, J., Ed.) American Chemical Society, Washington, DC (in press).
Kim, D., Yoshihara, K., & Ikegami, I. (1989) *Plant Cell Physiol. 30*, 679–684.
Kirsch, W., Seyer, P., & Hermann, R. G. (1986) *Curr. Genet. 10*, 843–855.
Lagoutte, B., & Mathis, P. (1989) *Photochem. Photobiol. 46*, 833–844.
Malkin, R. (1986) *FEBS Lett. 208*, 343–346.
Mathis, P., Ikegami, I., & Sêtif, P. (1988) *Photosynth. Res. 16*, 203–210.
Michel, H., Epp, O., & Deisenhoffer, J. (1986) *EMBO J. 5*, 2445–2451.
Nitschke, W., Feiler, U., Lockau, W., & Hauska, G. (1987) *FEBS Lett. 218*, 283–286.
Schoeder, H.-U., & Lockau, W. (1986) *FEBS Lett. 199*, 23–27.
Warncke, K., & Dutton, P. L. (1990) in *Current Research in Photosynthesis* (Baltscheffsky, M., Ed.) Vol. I, pp 157–160, Kluwer Academic Publishers, Dordrecht, The Netherlands.
Warncke, K., Gunner, M. R., Braun, B. S., Yu, C.-A., & Dutton, P. L. (1987) in *Progress in Photosynthesis Research* (Biggins, J., Ed.) Vol. 1, pp 225–227, Martinus Nijhoff Publishers, Dordrecht, The Netherlands.

Structure of the Membrane-Bound Protein Photosynthetic Reaction Center from *Rhodobacter sphaeroides*†

Chong-Hwan Chang,‡ Ossama El-Kabbani,§,∥ David Tiede,§ James Norris,§ and Marianne Schiffer*,‡
Biological and Medical Research Division and Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439
Received October 30, 1990; Revised Manuscript Received February 6, 1991

ABSTRACT: The structure of the photosynthetic reaction center (RC) from *Rhodobacter sphaeroides* was determined at 3.1-Å resolution by the molecular replacement method, using the *Rhodopseudomonas viridis* RC as the search structure. Atomic coordinates were refined with the difference Fourier method and restrained least-squares refinement techniques to a current *R* factor of 22%. The tertiary structure of the RC complex is stabilized by hydrophobic interactions between the L and M chains, by interactions of the pigments with each other and with the L and M chains, by residues from the L and M chains that coordinate to the Fe^{2+}, by salt bridges that are formed between the L and M chains and the H chain, and possibly by electrostatic forces between the ends of helices. The conserved residues at the N-termini of the L and M chains were identified as recognition sites for the H chain.

The photosynthetic reaction center (RC) is a transmembrane protein complex that carries out the light-induced charge separation that is the first step in photosynthesis. The structure of the RC from *Rhodopseudomonas viridis* was determined at 2.3-Å resolution (Deisenhofer & Michel, 1989). Chang et al. (1986) and Allen et al. (1986) used the coordinates of the *Rps. viridis* reaction center to solve the structure of the RC from *Rhodobacter sphaeroides* R-26 by the technique of molecular replacement. The RC from *Rb. sphaeroides* R-26 consists of three protein subunits, the L, M, and H chains. The L and M chains are transmembrane proteins with homologous structures and amino acid sequences. Each chain has five transmembrane helices, designated A, B, C, D, and E (Deisenhofer et al., 1985). The L and M chains are related to each other in the complex by a local 2-fold axis. Embedded in the L and M chains are nonprotein cofactors: four bacteriochlorophylls, two of which form the so-called "special pair" (BC_{LP} and BC_{MP}) and two that are monomeric "accessory" bacteriochlorophylls (BC_{LA} and BC_{MA}), two bacteriopheophytins (BP_L and BP_M), two quinones (Q_A and Q_B), and a

† C.-H.C. and M.S. were supported by the U.S. Department of Energy, Office of Health and Environmental Research, under Contract No. W-31-109-ENG-38 and by Public Health Service Grant GM36598; O.E.-K., D.T., and J.N. were supported by the U.S. Department of Energy, Division of Chemical Sciences, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38.
‡ Biological and Medical Research Division.
§ Chemistry Division.
∥ Present address: Center for Macromolecular Crystallography, University of Alabama at Birmingham, University Station, Birmingham, AL 35294.

0006-2960/91/0430-5352$02.50/0

non-heme iron atom. We are here reporting the structure and the general features of the RC from *Rb. sphaeroides* R-26 as determined at 3.1-Å resolution and refined by using restrained least-squares refinement methods.

Experimental Procedures

Purification of the Protein and Crystallization. The *Rb. sphaeroides* R-26 reaction center was isolated and purified as described earlier (Wraight, 1979); 1-octyl β-D-glucopyranoside was used as the detergent. The needle-shaped crystals were grown by vapor diffusion with poly(ethylene glycol) (PEG 4000) at pH 8 (Chang et al., 1985); some were up to 2 mm long.

Data Collection. The initial data to 3.7-Å resolution were collected by using the oscillation photographic technique with an Elliot GX20 rotating anode operating at 1.6 kW. The radiation used was Cu Kα with λ = 1.5418 Å, and the collimator was 0.3 mm. The Enraf-Nonius camera with eight cassettes was used; each casette contained three films. The crystal-to-film distance was 100.0 mm. The oscillation angle for each film was 2.0° with an exposure time of 4 h; there was no overlap in oscillation angle ranges between the exposures. After eight exposures, the crystal was translated along the spindle axis. The overall rotation range was 94° starting from a principal axis aligned along the direction of the X-ray beam. Throughout the data collection, only one crystal was used. The resulting films were scanned with an Optronics P1000 microdensitometer and processed by the method of Rossmann (1979). The R_{sym}[1] on intensities was 8.3% after the intensity was scaled by the Fourier–Bessel method (Weismann, 1979).

Higher resolution (3.1 Å) data were collected by using the synchrotron radiation source at Stanford Synchrotron Radiation Laboratory. The oscillation range was 2.5° with a 0.5° overlap between consecutive exposures. The exposure time was 2 min. After four exposures on each spot, the crystal was translated along the spindle axis. Two crystals were used for the data collection. For each crystal, there were eight possible exposure sites. The films were scanned at the Macromolecular Computational Facility at Purdue University. Each film was scanned and processed as described above, and the final results were obtained by postrefinement (Rossmann et al., 1979). The R_{sym} on intensities was improved by postrefinement from 11% to 9%. Although decay of intensities was kept to a minimum by exposing different parts of a long crystal, the quality of the data was affected by anisotropic mosaic spread, resulting in different spot shapes. Total number of reflections was 22866; among them, 13493 reflections (between 8 and 3.1 Å) were considered as observed with a $2.5\sigma F$ cutoff.

Molecular Replacement and Rigid Body Refinement. The structure of the *Rb. sphaeroides* RC was solved by the molecular replacement method, using the atomic coordinates of the *Rsp. viridis* RC and initial film data at 3.7-Å resolution (Chang et al., 1986). Several heavy metals were found to cause intensity differences. Unfortunately, these derivatives were not suitable for data collection because the heavy atoms also caused changes in the cell constants: the *b* axis became shorter by as much as 2%. The search for heavy atom derivatives was also hampered by the lack of suitable crystals.

Difference Fourier and Least-Squares Refinement. The positions of some missing side chains were located from a $(2F_o - F_c)$ difference map (Chang et al., 1986). The coordinates of the side chains were then included in the structure factor calculations for the next cycle of difference Fourier maps. After three cycles of this procedure, the positions of the side chains for 180 residues were located; the *R* factor decreased to 39%. After incorporation of the new data to 3.1-Å resolution, most missing side chains were located. At this stage, the structure model included 92% of the amino acid side chains of the RC molecule and the crystallographic *R* factor was 35.4% for the reflections between 8- and 3.1-Å resolution.

The structure of the RC from *Rb. sphaeroides* was further refined by using real-space refitting. Fragments comprising 30 or 50 sequential amino acid residues were omitted from the model. The remainder of the molecule was then used to calculate both a $2F_o - F_c$ electron density map and a SIGMAA map (Read, 1986) to minimize the bias resulting from the model.

The refinement of the RC was carried out with two different restrained-parameter least-squares refinement programs: PROLSQ (Hendrickson, 1985) and TNT (Tronrud et al., 1987). Since the standard group dictionary of PROLSQ does not include the coordinates for the pigment molecules, all atomic coordinates for the pigment molecules were held fixed during the PROLSQ refinement. An advantage of TNT was that it included the bacteriochlorophyll molecule in its standard group dictionary. We also added the coordinates for the ubiquinone molecule to this dictionary, which permitted the refinement of the protein and pigment atoms together. One cycle of TNT refinement, which uses fast Fourier transform algorithms, required about 2.3 central processor unit (cpu) hours on a MicroVax II for the 13493 reflections between 8- and 3.1-Å resolution and the approximately 7000 atoms of the RC molecule. The same computing task required 10 cpu minutes for PROLSQ refinement on the CRAY X/MP computer of the National Magnetic Fusion Energy Center at the Lawrence Livermore National Laboratory.

Electron density maps were displayed on an Evans & Sutherland PS300 computer graphics system, and the necessary adjustments were made with the program FRODO (Jones, 1978). The standard dictionary used in FRODO for the pigments was developed and given to us by Dr. J. Deisenhofer.

Identification of Helical Residues. For the determination of the last residue in a helical segment, the procedure described by Richardson and Richardson (1988) was followed. In this method, the residue "whose α-carbon is close to its correct spiral position on the helical cylinder" is considered part of the helix. The method is less sensitive to coordinate errors than one that uses hydrogen bonds or dihedral angles.

Results and Discussion

Progress of Refinement. The refinement procedure is summarized in Table I. In the early stages of model refinement, manual intervention was required. The omitted residues were adjusted, if necessary, and their side chains were removed if their electron densities were weak. In this case, these residues were replaced by alanines. As the refinement progressed, these side chains were returned if their corresponding electron density became interpretable. In order to improve the stereochemistry of the model, the restraint for nonbonded contacts was tightened from 0.5 to 0.3 Å during all PROLSQ refinement cycles that followed the second stage of model rebuilding and refinement.

At the end of each stage of model rebuilding, PROLSQ refinement was restarted by using the data between 8- and 3.1-Å resolution with an overall temperature factor. As the refinement approached convergence, isotropic refinement of individual temperature factors was allowed. The final mean temperature factor for all atoms in the RC structure model was 6 Å^2. This relatively low value is probably due to the

[1] $R_{sym} = \sum_{hkl}[|\overline{I_{hkl}} - I_{hkl}|]/\sum_{hkl} I_{hkl}$

Table I: Refinement of Data from 8- to 3.1-Å Resolution (13 493 Reflections)

Molecular Replacement Model
$2F_o - F_c$ map calculated
amino acid side chains that are missing from the model adjusted and fitted
R factor = 35.4%
calculate omit $2F_o - F_c$ and omit σ $2F_o - F_c$ maps
model refitted in omit maps after omitting 30 residues at a time
R factor = 33.7%
2 cycles of PROLSQ refinement
R factor = 30.2%
calculate omit $2F_o - F_c$ and omit σ $2F_o - F_c$ maps
model refitted in omit maps after omitting 50 residues at a time
R factor = 31.9%
21 cycles of PROLSQ refinement
R factor = 25.1%
calculate $2F_o - F_c$ map
model examined and manual adjustments made
3 cycles of geometry regularization (PROLSQ)
R factor = 31.7%
20 cycles of PROLSQ refinement
R factor = 24.5%
$2F_o - F_c$ map calculated
5 cycles of geometry regularization (TNT)
R factor = 25.3%
19 cycles of TNT refinement
R factor = 22.9%
amino acid residues that deviate from ideal geometry identified and refitted
R factor = 26.6%
17 cycles of PROLSQ
R factor = 23.3%
5 cycles of geometry regularization (TNT)
R factor = 26.6%
17 cycles of TNT refinement
R factor = 22.0%

Table II: Variation of *R* Factor with Resolution after PROLSQ Refinement

resolution (Å)	no. of reflections	observed[a] (%)	*R* factor[b]
8.0–6.5	1245	82	0.25
6.5–5.0	2887	75	0.24
5.0–4.0	4539	68	0.22
4.0–3.5	2868	43	0.23
3.5–3.2	1563	25	0.25
3.2–3.1	391	15	0.26

[a] With $F > 2.5\sigma F$; the *R* factor for all (21 126) reflections is 0.27. [b] *R* factor = $\sum||F_o| - |F_c||/\sum|F_o|$.

moderate resolution of the reflections used in the refinement. The temperature factors varied from 2 to 15 Å^2. The residues with the highest temperature factors were usually present on the surface of the molecule and belonged to either the periplasmic or cytoplasmic side of the complex. For TNT refinement, individual temperature factors were set at 9 Å^2. Individual temperature factors were not refined by TNT; when refinement of individual temperature factors was tried, the shifts in temperature factors were not reasonable.

During the course of the refinement process, the crystallographic *R* factor decreased from 35.4% to 22% for the 13 493 reflections between 8- and 3.1-Å resolution with $F > 2.5\sigma F$. This represented 59% of the observed reflections. We have looked at $2F_o - F_c$ electron density maps that were calculated with all reflections. These maps did not give additional information and were more noisy than the ones calculated by using reflections with $F > 2.5\sigma F$. In all, 60 cycles of PROLSQ least-squares refinement were performed. The root-mean-square (rms) shift of all atomic coordinates at the end of the TNT refinement was 0.01 Å, indicating convergence. The crystallographic *R* factor obtained from the PROLSQ refinement as a function of resolution is given in Table II. The deviations from ideal geometry at the end of the refinement, listed in Table III, demonstrate that the structure has good geometry. The average uncertainty in the atomic positions from a Luzzati plot (Luzzati, 1952), using all reflections, is 0.5 Å. The coordinates described here are deposited in the Brookhaven Protein Data Bank.

Table III: Deviations from Ideal Geometry at the End of Refinement

parameter	actual rms dev	target rms dev
PROLSQ rms Deviations from Ideal Values		
distance restraints (Å)		
bond distance	0.010	0.020
angle distance	0.027	0.030
planar 1–4 distance	0.033	0.050
planar restraint (Å)	0.008	0.020
chiral-center restraint (Å^3)	0.143	0.150
nonbonded contact restraints (Å)		
single torsion contact	0.263	0.300
multiple torsion contact	0.296	0.300
possible hydrogen bond	0.245	0.300
conformational torsion angle restraint (deg)		
planar (Ω)	1.5	3.0
TNT rms Deviations from Ideal Values		
bond distance (Å)	0.011	0.020
bond angle (deg)	2.7	3.0
planarity (trigonal) (Å)	0.008	0.020
planarity (other planes) (Å)	0.010	0.020
torsion angle[a] (deg)	20.2	15.0

[a] The torsion angles were not restrained during TNT refinement.

The Refined Model. The fit of the refined RC structure model to the electron density represents 97% of the amino acid side chains of the RC molecule. There is no electron density for the three C-terminal residues of the L chain and the two C-terminal residues of the M chain. In the H chain, the position of five residues at its C-terminus and 11 residues in the loop H45–H55 have not been determined. The electron density for parts of the phytyl tails of the bacteriochlorophylls and bacteriopheophytins are weak, suggesting high flexibility or static disorder for these segments. In general, the density for the tails is better on the L side: electron density is weak for CP15-CP17 of BC_{LA}, CP9-CP11 of BC_{MP}, CP6, CP10-CP12, CP16, and CP19-CP20 of BC_{MA}, and CP14-CP20 of BP_M. The electron density for the isoprenoid tails of Q_A and Q_B allowed the positioning of the first four isoprene units of the quinone tails.

An apparent anomaly is that electron densities for the Mg^{2+} atoms in the bacteriochlorophylls are very weak compared to those for the remainder of the molecule. This would suggest a partial occupancy of Mg^{2+} in the bacteriochlorophylls of the RC complex. However, following X-ray data collection, the crystals are still blue, and optical spectra recorded with single crystals redissolved following X-ray data collection are indistinguishable from those recorded before crystallization. These observations suggest that Mg^{2+} was present in the bacteriochlorophylls. We do not yet understand this anomaly. Among four BCLs, the Mg^{2+} in BC_{MA} has the strongest electron density but nevertheless appears to have less than full occupancy. On the other hand, in the structure of the RC derived from the wild-type Y-strain *Rps. sphaeroides*, unambiguous electron density is present for all of the Mg^{2+} atoms. The Y-strain RC structure was solved by molecular replacement using the R-26 structure described here (Arnoux et al., 1989). The refinement of the Y-strain RC structure is in progress.

Of a total of 827 α-carbon atoms in the final model, 129 shifted more than 1 Å during model rebuilding and refinement. α-Carbon atoms of the L or M subunit that shifted more than 1 Å are mainly located in extended strands connecting the membrane-spanning helices and are present in periplasmic or

cytoplasmic regions of the complex. Most α-carbon atoms of the H subunit that shifted more than 1 Å are located in strands forming surface loops, which are present in the cytoplasmic side of the complex. α-Carbon atoms located next to amino acid residues that were inserted or deleted from the initial model also shifted more than 1 Å during the refinement process.

Description of the Complex. As the success of the molecular replacement method suggests, the overall structure of the RC from *Rb. sphaeroides* is very similar to that of *Rps. viridis* (Deisenhofer & Michel, 1989) and to the structure of *Rb. sphaeroides* that was determined independently [Komiya et al. (1988) and the rest of the series thereon], also by using the *Rps. viridis* structure. Therefore, in this paper, the description of the general common features of the structure will not be repeated in detail. The M-chain residues in the *Rb. sphaeroides* RC are labeled according to the *Rps. viridis* RC following the sequence alignment described by Michel et al. (1986).

The RC from *Rb. sphaeroides* is a complex formed from three protein subunits, the L, M, and H chains, and eight chromophores. There are two bacteriochlorophyll molecules that form the special pair, two accessory bacteriochlorophylls, two bacteriopheophytins, and two ubiquinones. The complex also contains a non-heme Fe^{2+} atom. The L and M chains are transmembrane proteins with homologous structures. Each chain contains five transmembrane helices; these transmembrane helices are the most striking feature of the structure. A local, noncrystallographic, 2-fold axis relates the M and L subunits. The H chain has only one transmembrane helical segment; most of the H subunit is located on the cytoplasmic side of the membrane. The local 2-fold axis of the complex does not apply to the H chain. The complex can function in vitro without the H chain, though with diminished efficiency (Debus et al., 1985). The H chain does not participate directly in the accommodation of the chromophores, but it increases the barrier between Fe^{2+} and quinones with respect to the cytoplasm. The H chain is considered to be important in the proper assembly of the complex in the membrane (Chory et al., 1984). Accordingly, it can be speculated that the role of the single membrane-spanning helix of the H subunit is related to maintaining the RC complex in the proper functional orientation in the membrane.

The tertiary structure of the RC complex is stabilized by hydrophobic interactions between the L and M chains, by interactions of the pigments with each other and with the L and M chains, by residues from the L and M chains that coordinate to the Fe^{2+}, by salt bridges that are formed between the L and M chains and the H chain, and possibly by electrostatic forces between the ends of helices. The topology of the complex is such that the L and M chains, the four bacteriochlorophylls, and the two bacteriopheophytins are required together for its assembly.

The buried surface area (Lee & Richards, 1971) between the L and M chains is 3990 Å^2, that between the H and L chains is 1870 Å^2, and that between the H and M chains is 2770 Å^2. Although the bacteriochlorophylls of the special pair are mostly in contact with one of the chains, the monomeric bacteriochlorophylls and the bacteriopheophytins are in contact with both chains. Figure 1 illustrates the locations of the cofactors in the complex. The details of the pigment protein contacts are discussed in the following paper (El-Kabbani et al., 1991).

The interaction between the L and M chains is also facilitated by hydrogen bonds; some of these are near the periplasmic surface and some are in the transmembrane region. Near the periplasmic surface, a hydrogen bond is formed between asparagines M185 and L166. Nearby, conserved residues Asp M182 and Tyr L169 form a hydrogen bond, although the symmetry-related residues Asp L155 and Tyr M196 do not form hydrogen bonds. Salt bridges and their dimensions are summarized in Table IV. Approximately one-third of the salt bridges are interchain bridges between the H subunit and the M and L subunits.

Table IV: List of Possible Salt Bridges

residues		d (Å)	residues		d (Å)
Intrachain			Interchain		
Arg L10	Glu L6	3.7	Arg L7	Glu H43	4.8
Lys L110	Asp L23	2.7[a]	Lys L8	Glu H81	2.9
Arg L207	Glu L205	4.6	Arg M231	Glu H122	2.5
Arg L217	Asp L213	4.8	Arg M231	Glu H230	3.0
Arg M86	Asp M87	3.9	Arg M239	Glu H38	2.8
Lys M108	Glu M109	4.9	Arg M239	Glu H79	4.6
Arg M162	Glu M171	2.9[a]	Arg M265	Glu H34	4.7
His M217	Glu M232	4.1[a]	Arg H117	Glu M234	3.6
Arg M239	Glu M244	4.4	Arg H118	Glu M234	4.8
His M264	Glu M232	2.7[a]	Arg H177	Glu M230	3.6
Arg M265	Glu M261	4.6	Lys L8	Glu M244	4.2
Arg H37	Glu H34	2.9	His L116	Glu M2	3.0
Lys H106	Asp H103	4.7	His L190	Glu M232	3.7
Arg H117	Asp H231	3.2[a]	Arg L217	Asp M17	2.2
Lys H130	Asp H170	4.1[a]	His L230	Glu M232	3.6
Lys H130	Glu H173	4.2[a]			
Lys H135	Glu H180	4.3[a]			
Arg H177	Asp H170	2.5[a]			
Arg H202	Glu H159	3.2[a]			

[a] Intrachain salt bridges that connect chain segments that are separated by more than four residues.

Parts of the chromophores are in contact with the detergent in the crystal or with the lipids in the membrane. The most exposed chromophores, in which about 10% of the macrocycles are in contact with the membrane, are the monomeric bacteriochlorophylls. The bacteriochlorophylls of the special pair are completely buried with the exception of the last few atoms of the phytyl tails. The phytyl tails of the monomeric bacteriochlorophylls and the bacteriopheophytins are in contact with the membrane. The phytyl tails effectively shield the bacteriopheophytin and the bacteriochlorophyll of the special pair macrocycles from the membrane. The tail of BC_{LA} is partially shielded by the transmembrane helix of the H chain. In *Rps. viridis*, this segment has a different configuration: it folds over its own macrocycle (Deisenhofer & Michel, 1989). Of course, it is not clear if the conformation of the tails observed in the crystal is the only conformation that can be adopted in the membrane.

The orientation of the helices with respect to each other in the L and M chains is shown schematically in Figure 2. They tend to be antiparallel and make an angle of 25° or less with each other. Within one subunit, only the C and E helices are parallel to each other. Upon formation of the complex, the D and E helices of the L and M subunits form a four-helix bundle, an important packing feature of helices previously recognized in other proteins (Weber & Salemme, 1980; Sheridan et al., 1982). The antiparallel arrangement of helices within each chain and in the complex might add to the stabilization of the molecule. Table V gives the electrostatic stabilization of the complex due to the dipoles of the transmembrane helices (Hol et al., 1978). The calculated stabilization energies are of the same order of magnitude as those reported by Sheridan et al. (1982).

Transmembrane Helices of the L and M Chains. The membrane-spanning helices from A to E are L32–L55, L84–

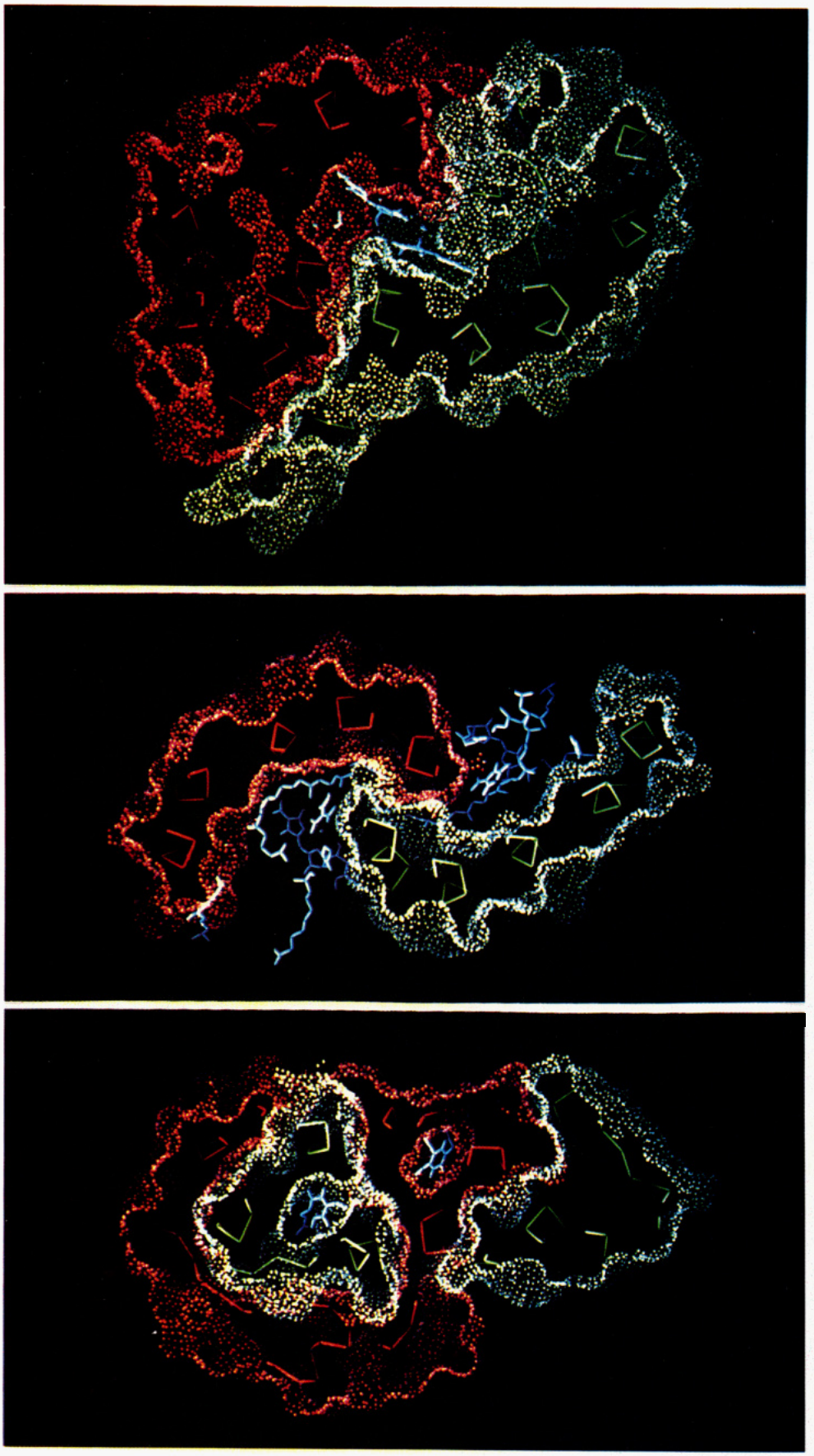

FIGURE 1: Sections through the reaction center molecule perpendicular to the local 2-fold axis. The sections are viewed from the cytoplasmic side of the membrane. The surface and backbone of the L chain are shown in green. The surface and backbone of the M chain are shown in red. The cofactors are shown in blue. The H chain is not shown. (Top panel) Illustration of how the special pair is buried within the complex by the L and M chains. (Middle panel) The accessory chlorophylls and the pheophytins are located between the L and M chains and are also exposed to the solvent or the membrane. The pigments on the two sides are separated by the protein near the middle of the complex. This figure shows the cross section of the four-helix bundle formed by the D and E helices from the two chains. (Bottom panel) The asymmetry of the chain distribution around the level of the quinones is illustrated.

Table V: Electrostatic Stabilization (kcal/mol) Due to Transmembrane Helices[a]

E_L	−4.2
E_M	−4.2
E_{L+M}	−9.6
energy gain on L + M association	−1.2 (13% of total)
E_{L+M+H}	−11.3
energy gain on association of H	−1.7 (15% of total)

[a] Assuming $\pm 0.5E$ at the end of each helix located on the helix axis; dielectric constant = 4 (Hol et al., 1978; Sheridan et al., 1982).

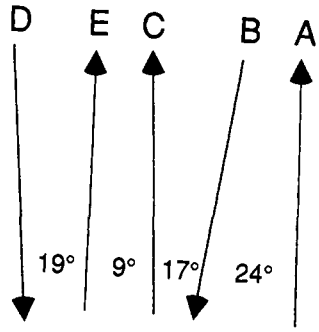

FIGURE 2: Schematic illustration of chain directions in the transmembrane helices of the L or M chain. The angles between neighboring helices are shown.

L111, L115–L140, L170–L199, and L225–L251 and M52–M77, M101–M139, M142–M167, M197–M224, and M259–M285 in the L and M subunits and H13–H40 in the H subunit. The designation of the residues in the helix is based on the method described by Richardson and Richardson (1988) and may be slightly different from others (Deisenhofer & Michel, 1989; Allen et al., 1987). There are two helical regions on the periplasmic surface of both subunits. One of these regions comprises helices (CD) formed by residues L155–L163 and M179–M188. Because of their location, these are amphipathic helices that have a well-defined hydrophobic side and a well-defined polar side. The polar side points toward the periplasm and can form the surface of contact with the soluble cytochrome c_2 molecule. The other helical region is located near the C-terminus of both chains, L259–L267 and M292–M299. The M chain has an additional short helix, residues M81–M86 after its A helix. Allen et al. (1987) also identified residues L59–L64 as helical. In our structure, the chain segment between the A and B helices of the L chain forms an irregular loop at L59–L64. These residues (L59–L64) were also assigned to a nonhelical region in *Rps. viridis* by Deisenhofer and Michel (1989). On the cytoplasmic side of the complex, helical segments in both subunits, L210–L220 and M241–M251, form part of the quinone binding sites. The M subunit has an additional short helical segment, M232–M237, located on the cytoplasmic side of the Fe^{2+} atom.

The transmembrane helices of the RC are characterized by the absence of charged residues in the middle of the helical regions. In each helix, there exists a segment of at least 16 residues uninterrupted by Glu, Asp, Lys, Arg, or His residues. No other segment of the RC molecule has this characteristic. Therefore, this unusual feature, shared by the primary structures of all transmembrane helices of the reaction center complexes, suggests the possibility that corresponding segments of uncharged amino acids in other membrane-associated proteins might reveal the location of transmembrane helices. It is of interest that there are some polar residues in the membrane-spanning helices. Most of these side chains form hydrogen bonds with either protein constituents or the chromophores. Only three of the 31 polar L-chain residues and 12 of the 45 polar M-chain residues fail to form hydrogen bonds. There are three interhelix hydrogen bonds formed by side chains: Arg L135–Thr L251 between the C and E helices of the L chain, Ser M150–Thr M275 between the C and E helices of the M chain, and Trp M266–Asn H35 between the E helix of the M chain and the transmembrane helix of the H chain. The distances vary from 2.6 to 3.3 Å.

There are kinks in helices C and E of both the L and M chains. The presence of a proline at L136/M163 in the C helices causes the carbonyl atom of Leu L131/M158 to form a hydrogen bond other than an intrahelix hydrogen bond. Consequently, there is a marginal deviation in the backbone from the regular helix conformation. In the case of the E helix, it is not a proline but the special pair bacteriochlorophylls near the middle of the helix that cause the helices to kink away from the special pair (Figure 3). Proline L118, which is located at the start of the C helix, does not cause a kink in the helix because the backbone nitrogen atom is not required to form a hydrogen bond since it is on the first turn of the helix. Proline L118 might be responsible for the difference in conformation at the N-terminal part of the C helix in the L and M subunits.

Sections That Differ between the L and M Chains. The 2-fold symmetry between the L and M subunits is broken near the cytoplasmic side of the complex. The N-terminal segment of the M chain (before the A helix) has 20 more amino acids than that of the L chain. Further, there is an additional insertion of seven residues in the M chain following the D helix that is important for the function of the complex.

The 31 amino acid segment at the N-terminus of the L chain, before the A helix, is of interest. This sequence is

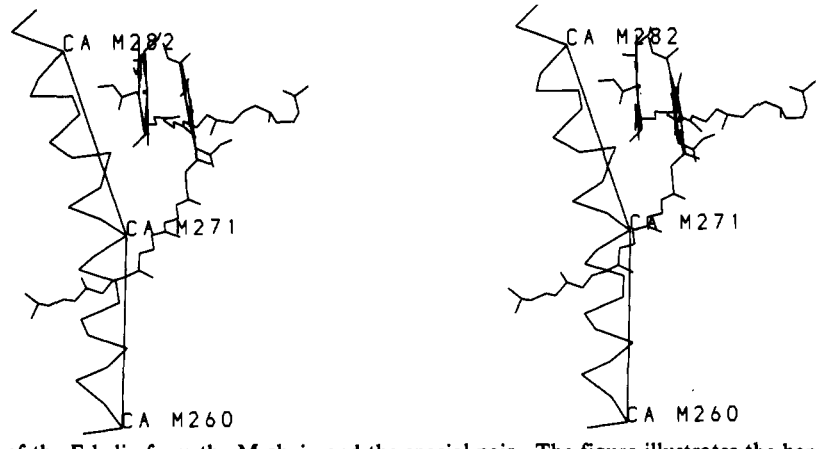

FIGURE 3: Stereo diagram of the E helix from the M chain and the special pair. The figure illustrates the bend in the E helix, apparently to accommodate the special pair.

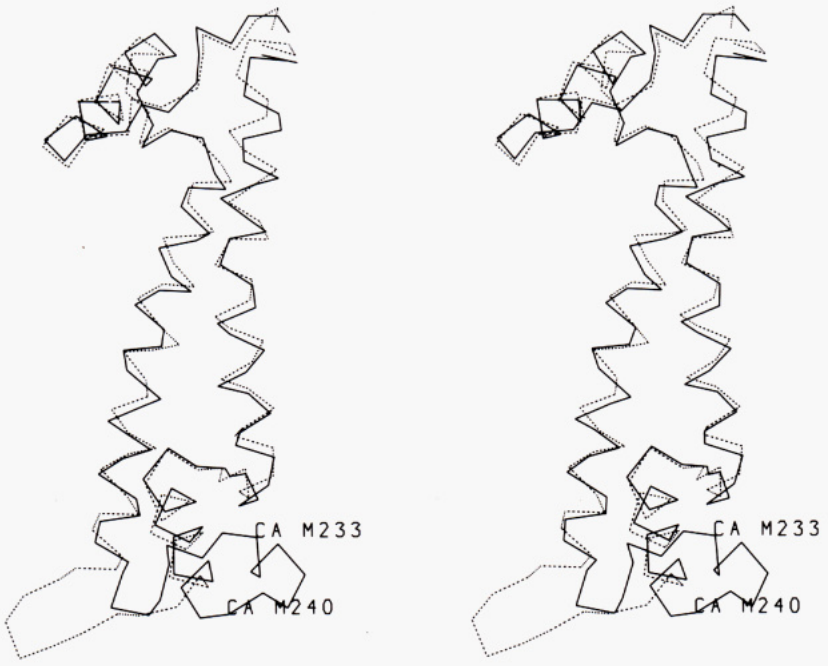

FIGURE 4: Stereo diagram that shows the loops connecting the D and DE helices in the L and M chains. For easier comparison, the L and M chains were overlapped; the M chain is drawn in a solid line, and the L chain is drawn in a dotted line.

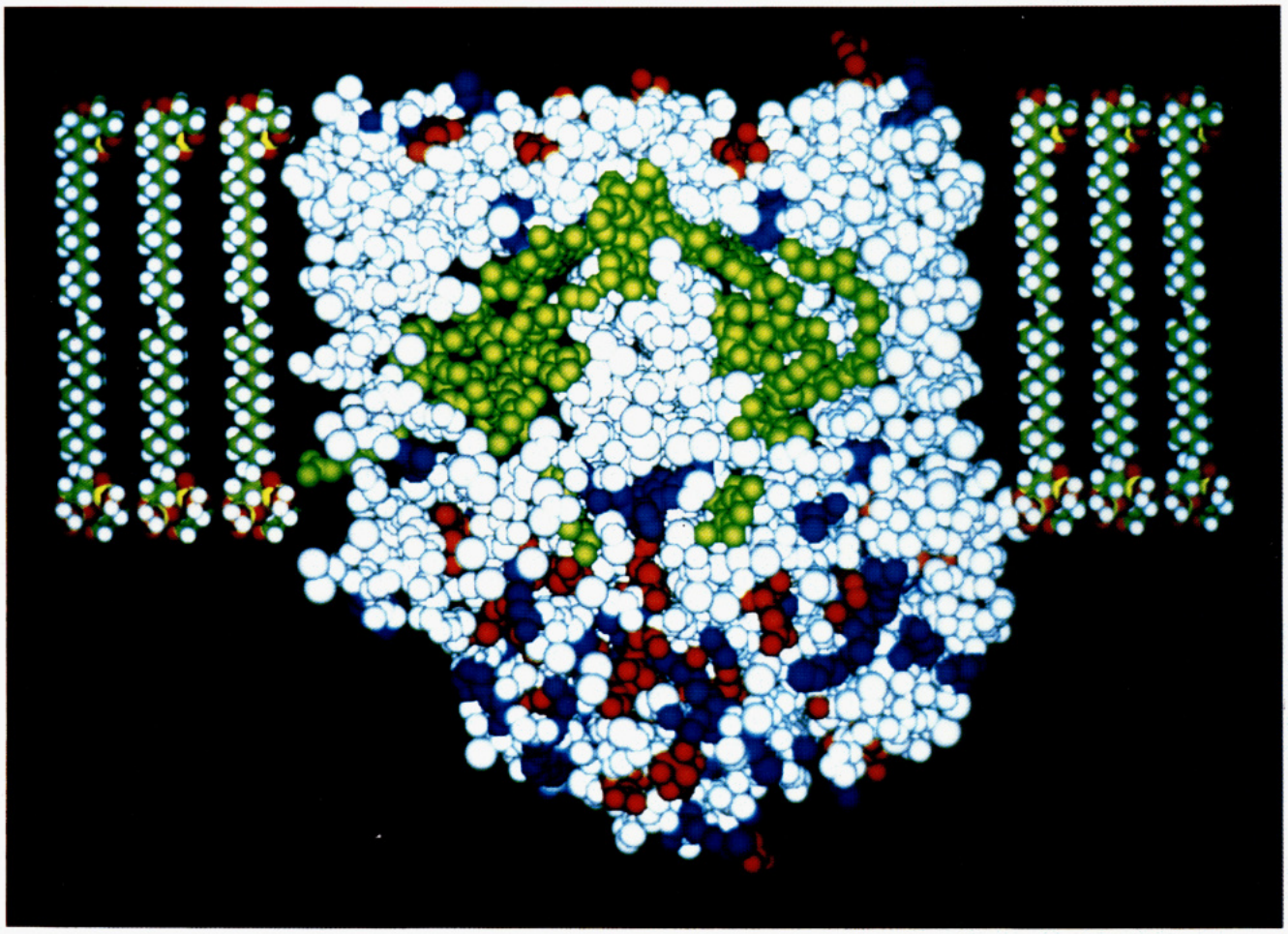

FIGURE 5: A space-filling model of the cross section of the reaction center is shown embedded in a computer-generated model membrane. The atoms of the cofactors are shown in green. Atoms of the protein are shown in white, except that atoms of acidic residues are red and atoms of basic residues are blue. The periplasmic side of the membrane is at the top of the figure, and the cytoplasmic side of the membrane is at the bottom of the figure. Most of the H chain is located at the cytoplasmic side of the membrane.

conserved among the different species. This is unusual because in most proteins the amino acid sequences of the N- and C-terminal ends form loose looplike structures and are not conserved. Among the 31 residues, only five are different among four purple photosynthetic bacteria: *Rb. sphaeroides* (Williams et al., 1983, 1984), *Rhodobacter capsulatus*

Table VI: Residues Conserved in the L and M Subunits (260 Residues Total) of *Rps. viridis*, *Rb. spheroides*, and *Rb. capsulatus*

	of all residues	of residues in transmembrane helices
Gly	14	7
Pro	8	2
Trp	6	2
His	5	4
Leu	4	2
Ile	3	2
Arg	3	2
Phe	1	1
Tyr	1	0
Asp	1	0
Asn	1	0
total	47 (18%)	22 (16%)

(Youvan et al., 1984), *Rps. viridis* (Michel et al., 1986), and *Rhodospirillum rubrum* (Belanger et al., 1988). However, if changes from tyrosine to phenylalanine and isoleucine to valine are considered as conservative substitutions, then only two residues are functionally altered, arginine to proline at L12 and aspartate to asparagine or serine at L20. This homology is not conserved in the green bacterium *Chloroflexus aurantiacus* (Ovchinnikov et al., 1988a), which has 34 more amino acids at the N-terminus of its L chain. These extra amino acids in the *Chl. aurantiacus* RC might be correlated with the fact that this RC is not believed to have an H chain (Pierson et al., 1983).

The M chain has 51 amino acids from the N-terminus to the A helix. In contrast to the L chain, little sequence homology is observed for the 43 N-terminal amino acids. The eight amino acids before the A helix have a good sequence homology among all five species, including *Chl. aurantiacus* (Ovchinnikov et al., 1988b). The equivalent L site has also conserved residues among five species. Residues Gly L27 and Pro L28 form a turn after a straight segment. Although M47 Gly and M48 Pro are conserved and in the homologous position relative to those of the L chain, they do not form a turn but are in the middle of a straight segment.

The cytoplasmic domain of the H chain does not conform to the local 2-fold axis of the RC. It is asymmetrically located; therefore, it contacts different segments of the L and M chain. The L-chain segment that includes residues 1–26 is in contact with the H chain; 23 of these residues are conserved. The N-terminal residues of the M chain are less conserved. The segment that consists of the first 15 residues is in contact with the H chain, but these contacts are less extensive than those of the L chain. The conservation of amino acid sequences in the N-terminal segment of the L and M chains is important for the recognition and interaction with the H chain.

Another interesting feature occurs at the loop connecting the D and DE helices. In Figure 4, the L and M chains are overlapped to illustrate the difference in these segments. There are 11 residues in the L chain, L199–L208, and 18 residues in the M chain, M225–M242. The loop formed by the M chain is located on the cytoplasmic side of the non-heme iron and contains Glu M232, which binds the iron. Residues M232–M238 are in a helical conformation. Four residues (M230, M231, M234, and M239) of this loop form salt bridges with residues from the H chain (Table V). All but one of the M-chain residues that form salt bridges with the H-chain residues are from this loop. The equivalent loop of the L chain turns away from the core of the RC; it points in the plane of the membrane and is inserted between the loop connecting the B and C helices in the M chain and the sheet of the H chain, near H45. The presence of the fifth ligand to the Fe^{2+}, Glu M232, and the central location of this loop suggest that this feature may play an important role in the function of the RC. This chain segment and the chain at the N-terminus represent the major features that do not follow the symmetry between the L and M chains.

Functions of the Conserved Residues of the L and M Chains. The architectures of the L and M subunits are strikingly similar; this is reflected in the conservation of many amino acid residues. In six chains (the L and M chains of three RCs from the closely related organisms *Rps. viridis*, *Rb. sphaeroides*, and *Rb. capsulatus*), 47 residues are identical. These conserved residues are probably required to maintain the structure of the subunits and their interactions with the cofactors (Michel & Deisenhofer, 1988; Komiya et al., 1988). Remarkable is the conservation of Gly, Pro, and Trp residues (Table VI). The His residues coordinate to the bacteriochlorophylls and to the iron atom, while the conserved Gly residues are near the beginnings and ends of helical segments. The Pro residues are predominantly in nonhelical segments. Trp residues might have special functions in transmembrane proteins using both the hydrogen-bonding capabilities and the hydrophobic character of this residue.

The three-dimensional structures of the L and M chains are homologous, but their function in electron transfer differs. Although a local 2-fold axis relates the L and M chains in the RC, light-induced electrons flow exclusively through the pigments associated with the L chain (Michel-Beyerle et al., 1988). To understand the origins of the different functions of the L and M chains, positions where the two chains have different amino acids were examined. At 29 positions, L and M chains differ, but the residue within each chain type is conserved (Table VII). Some of these residues are involved in H-chain contacts that are not symmetric; others are candidates in determining the asymmetry of the electron flow in the complex (discussed in the following article). The reasons for the differences in many of the residues are not understood.

Location of the Reaction Center in the Lipid Bilayer. Two principal evolutionary "forces" combined to construct the RC architecture. The energy-transducing role of this protein complex defined the composition, location, and microenvironmental chemistries of the chromophores, as well as other constituents of the electron transport system. The requirement that the electron be transported across a barrier requires the RC to span the barrier and to possess surface properties compatible with multiple environments.

The RC is a transmembrane protein that is designed to pass through and function in three different environments. There are two distinct environments within the membrane bilayer itself. The center portion of the bilayer, about 30 Å in width, is hydrophobic or lipophilic. The two surfaces of the bilayer are formed from polar lipid head-groups of approximately 5 Å in width. Outside the membrane in the cytoplasm or in the periplasm, the environment is polar. On the basis of alignment with a model membrane (Figure 5), the L and M subunits are flush with the membrane surface on the periplasmic side. The bulk of the H subunit is located in the cytoplasm. The quinone head-groups are at the level of the membrane head-groups, and the bacteriochlorophylls and bacteriopheophytins are at the level of the hydrophobic part of the bilayer. While the cytoplasmic portion of the H subunit has many polar and charged residues, the charged residues in the L and M chains are distributed near the lipid head-group of the membrane surfaces. Most of the interactions of the L and M chains with each other and with the pigments are nonpolar, but the interfaces between the H chain and the L and M chains have several salt bridges.

Table VII: Residues That Are Conserved in the L Chain and Conserved but Different in the M Chain of RCs of *Rps. viridis*, *Rb. sphaeroides*, and *Rb. capsulatus*[a]

F	L22[b]	G	M42[b]	L	L193[e]	Thr	M220[e]
W	L25[b]	Gln	M45[b]	Ser	L196	A	M223
V	L31[b]	L	M51[b]	Asn	L199	Arg	M226
I	L89	A	M116	Thr	L208	A	M242
C	L92	F	M119	**His**	L211	Arg	M245
A	L93	M	M120	Glu	L212[e]	A	M246[e]
V	L105	**Tyr**	M132	F	L216[e]	W	M250[e]
I	L107	Arg	M134	Asp	L218	W	M256[e]
P	L108	A	M145	**Tyr**	L222[e]	F	M256[e]
G	L143	Ser	M170	Ser	L223[e]	Asn	M257[e]
G	L161[c]	Ser	M188[c]	G	L225	Thr	M259[b]
M	L174[d]	G	M201[d]	L	L227	Glu	M261[b]
F	L181[c,d]	**Tyr**	M208[c,d]	L	L232	W	M266[b]
A	L186[e]	L	M213[e]	W	L266[b]	**His**	M299[b]
L	L187	F	M214				

[a] Nonpolar residues are listed by their one-letter codes, aromatic residues are listed in boldface type, and ionizable or polar residues are listed by their three-letter codes. [b] Near H chain. [c] Near special pair. [d] Near accessory bacteriochlorophylls. [e] Near quinones.

RCs from two bacterial species represent the first transmembrane proteins for which the detailed three-dimensional structures are known. These structures serve as models for deriving rules for membrane proteins. Some of the observations made on the RC structure are relevant to membrane proteins in general; others are related to the specific function of the complex. The RC structure in *Rb. capsulatus* and *Rb. sphaeroides* can be changed by genetic engineering techniques; therefore, the effect of individual residues can be explored.

ACKNOWLEDGMENTS

We thank Dr. F. J. Stevens and Dr. D. K. Hanson for many useful suggestions and for reviewing the manuscript.

REFERENCES

Allen, J. P., Feher, G., Yeates, T. O., Rees, D. C., Deisenhofer, J., Michel, H., & Huber, R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8589–8593.

Allen, J. P., Feher, G., Yeates, T. O., Komiya, H., & Rees, D. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6162–6166.

Arnoux, B., Ducruix, A., Reiss-Husson, F., Lutz, M., Norris, J., Schiffer, M., & Chang, C.-H. (1989) *FEBS Lett. 258*, 47–50.

Belanger, G., Berard, J., Corriveau, P., & Gingras, G. (1988) *J. Biol. Chem. 263*, 7632–7638.

Chang, C.-H., Schiffer, M., Tiede, D., Smith, U., & Norris, J. (1985) *J. Mol. Biol. 186*, 201–203.

Chang, C.-H., Tiede, D., Tang, J., Smith, U., Norris, J., & Schiffer, M. (1986) *FEBS Lett. 205*, 82–86.

Chory, J. T., Donohue, J., Varga, A. R., Staehelin, L. A., & Kaplan, S. (1984) *J. Bacteriol. 159*, 540–554.

Debus, R. J., Feher, G., & Okamura, M. Y. (1985) *Biochemistry 24*, 2488–2500.

Deisenhofer, J., & Michel, H. (1989) *EMBO J. 8*, 2149–2169.

Deisenhofer, J., Epp, O., Miki, R., Huber, R., & Michel, H. (1985) *Nature 318*, 618–624.

El-Kabbani, O., Chang, C.-H., Tiede, D., Norris, J., & Schiffer, M. (1991) *Biochemistry* (following paper in this issue).

Hendrickson, W. (1985) *Methods Enzymol. 115*, 252–270.

Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. (1978) *Nature 273*, 443–446.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Komiya, H., Yeates, T. O., Rees, D. C., Allen, J. D., & Feher, G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9012–9016.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.

Michel, H., & Deisenhofer, J. (1988) *Biochemistry 27*, 1–7.

Michel, H., Weyer, K. A., Gruenberg, H., Dunger, I., Oesterhelt, D., & Lottspeich, F. (1986) *EMBO J. 5*, 1149–1158.

Michel-Beyerle, M. E., Plato, M., Deisenhofer, J., Michel, H., Bixon, M., & Jortner, J. (1988) *Biochim. Biophys. Acta 932*, 52–70.

Ovchinnikov, Yu A., Abdulaev, N. G., Zolotarev, A. S., Shmuckler, B. E., Zargarov, A. A., Kutuzov, M. A., Telezhinskaya, I. N., & Levina, N. B. (1988a) *FEBS Lett. 231*, 237–242.

Ovchinnikov, Yu A., Abdulaev, N. G., Shmuckler, B. E., Zargarov, A. A., Kutuzov, M. A., Telezhinskaya, I. N., Levina, N. B., & Zolotarev, A. S. (1988b) *FEBS Lett. 232*, 364–368.

Pierson, B. K., Thornber, J. P., & Seftor, R. E. B. (1983) *Biochim. Biophys. Acta 723*, 322–326.

Read, R. (1986) *Acta Crystallogr., Sect. A 42*, 140–149.

Richardson, J., & Richardson, D. (1988) *Science 240*, 1648–1652.

Rossmann, M. (1979) *J. Appl. Crystallogr. 12*, 225–238.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meduid, S. S., & Tsukihara, T. (1979) *J. Appl. Crystallogr. 12*, 570–581.

Sheridan, R. P., Levy, R. M., & Salemme, F. R. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4545–4549.

Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr., Sect. A 43*, 489–501.

Weber, P. C., & Salemme, F. R. (1980) *Nature 287*, 82–84.

Weissmann, L. (1979) Ph.D. Thesis, University of California at Los Angeles.

Williams, J. C., Steiner, L. A., Ogden, R. C., Simon, M. I., & Feher, G. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6505–6509.

Williams, J. C., Steiner, L. A., Feher, G., & Simon, M. I. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7303–7307.

Wraight, C. A. (1979) *Biochim. Biophys. Acta 548*, 309–327.

Youvan, D. C., Bylina, E. J., Alberti, M., Bergusch, H., & Hearst, J. E. (1984) *Cell 37*, 947–957.