The Solution Structure of *Rhodobacter sphaeroides* LH1β Reveals Two Helical Domains Separated by a More Flexible Region: Structural Consequences for the LH1 Complex

Matthew J. Conroy¹, Willem H. J. Westerhuis¹
Pamela S. Parkes-Loach², Paul A. Loach², C. Neil Hunter¹ and Michael P. Williamson¹*

1Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK
2Department of Biochemistry Molecular Biology and Cell Biology, Northwestern University, Evanston IL 60208-3500, USA

Here, the solution structure of the *Rhodobacter sphaeroides* core light-harvesting complex β polypeptide solubilised in chloroform:methanol is presented. The structure, determined by homonuclear NMR spectroscopy and distance geometry, comprises two alpha helical regions (residue −34 to −15 and −11 to +6, using the numbering system in which the conserved histidine residue is numbered zero) joined by a more flexible four amino acid residue linker. The C-terminal helix forms the membrane spanning region in the intact LH1 complex, whilst the N-terminal helix must lie in the lipid head groups or in the cytoplasm, and form the basis of interaction with the α polypeptide. The structure of a mutant β polypeptide Wβ9F was also determined. This mutant, which is deficient in a hydrogen bond donor to the bacteriochlorophyll, showed an identical structure to the wild-type, implying that observed differences in interaction with other LH1 polypeptides must arise from cofactor binding. Using these structures we propose a modification to existing models of the intact LH1 complex by replacing the continuous helix of the β polypeptide with two helices, one of which lies at an acute angle to the membrane plane. We suggest that a key difference between LH1 and LH2 is that the β subunit is more bent in LH1. This modification puts the N terminus of LH1β close to the reaction centre H subunit, and provides a rationale for the different ring sizes of LH1 and LH2 complexes.

Keywords: light-harvesting complex; transmembrane helix; NMR; *Rhodobacter sphaeroides*; model

*Corresponding author

Introduction

Purple photosynthetic bacteria contain two types of light-harvesting or antenna complex which are responsible for the efficient transfer of light energy to the reaction centre (RC), the site of charge separation. The core antenna complex, LH1, surrounds the RC (Walz *et al.*, 1998) and exists in a fixed stoichiometry with it. The peripheral complex, LH2, is not as intimately associated with the RC and the amount of this complex varies in response to environmental changes.

Both complexes are made up of multiple copies of two integral membrane polypeptides, termed α and β, to which bacteriochlorophyll (Bchl) and carotenoid are bound. In general, these polypeptides are approximately 50 residues long and are highly conserved in sequence between species and between the different LH complexes (Brunisholz & Zuber, 1992). LH1 binds one Bchl species absorbing at around 875 nm (termed B875), whereas LH2 binds two Bchl species, B800 and B850. The crystal
structures of two LH2 complexes have been solved (McDermott et al., 1995; Koepke et al., 1996) and projection maps of LH1 complexes are available (Walz et al., 1998; Karrasch et al., 1995) which show that the LH1 complex is probably similar in architecture to LH2, although somewhat larger.

To date, no atomic resolution structure of LH1 is available although several models exist based on LH2 structures and LH1 projection maps (Hu et al., 1998; Hu & Schulten, 1998; Papiz et al., 1996) and on sequence comparisons (Olsen & Hunter, 1994) in conjunction with the available data on pigment-protein interactions. Whilst the gross architecture of LH1 is likely to be similar to that of LH2, it will undoubtedly differ in detail. In particular, LH1 complexes have no B800 Bchl and thus have no requirement to form this binding site. The LH1 complex must also interact with the RC and with LH2.

Comparatively little is known about membrane protein:co-factor interactions and, in particular, little is known about whether co-factor binding induces conformational change in the protein, or whether the protein forms a rigid pocket into which the co-factor fits. One aim of this study was to investigate the structure of the protein in the absence of pigment co-factors to permit comparisons with related intact complexes.

The LH1β polypeptide is of particular interest for several reasons: it lies on the outside of the complex (by analogy with LH2) and may therefore interact with LH2; its N terminus is likely to point towards the centre of the complex and may also interact with the RC as well as with LH1x; and β is also able to form a B820 subunit complex not just with α (βBchlα) but with another β polypeptide and Bchl (βBchlβ) (Parkes-Loach et al., 1988; Loach et al., 1994).

Here, we present the solution structure of the LH1β polypeptide from *Rb. sphaeroides* in organic solvent along with that of a mutant deficient in a hydrogen bond to the Bchl co-factor termed W_8_9F. Our preliminary work on the wild-type polypeptide (Kikuchi et al., 1999) illustrated its helical nature and also showed that the structure in organic solvents is the same as in detergent micelles. Previous studies of membrane proteins in the solvent system used here (Pervushin et al., 1994 and references therein; Girvin & Fillingame, 1993, 1994; Schweiger et al., 1998) and in similar solvents (Rastogi & Girvin, 1999) have illustrated that structural integrity is maintained in spite of the protein being removed from its native environment. The structural consequences for the bend in LH1β have been explored and a model of the LH1 complex is proposed.

**Results**

**Assignment**

In order to obtain high quality NMR spectra of LH1β in methanol:chloroform (Figure 1) it was found to be necessary to pre-treat with trifluoroacetic acid, as described in Materials and Methods. Although spectra were significantly overlapped, assignment using standard homonuclear techniques (Basus, 1989; Wüthrich, 1986) was possible. All backbone signals could be assigned with the exception of Ala_37 (using the standard (Olsen & Hunter, 1994; Brunisholz & Zuber, 1992) numbering protocol in which the conserved histidine ligand to the Bchl is numbered zero: Ala_–37 is thus the N-terminal residue). Almost all side-chain resonances were assigned. Chemical shift assignments have been deposited with BioMagResBank (accession number 4303).

**Secondary structure prediction**

The secondary shift (i.e. the difference between the observed shift and the expected random coil value) of all alpha protons was calculated relative to the random coil data (Wishart et al., 1992). Some large negative secondary shifts were observed which imply a helical structure (Wishart et al., 1992). The chemical shift index structure prediction (Wishart & Sykes, 1994) indicated that the protein has two helical regions (residues -26 to -13 and -7 to +4), and further indicated that there is a five-residue break in the helical structure between the two helices. Both termini of the polypeptide (residues -37 to -27 and +5 to +10) were also predicted to be unstructured. CSI data are summarised in Figure 2.
A total of 442 structurally useful NOES were unambiguously identifiable for the wild-type poly-peptide and were included in the structure calculation. Of these, 164 were sequential and 136 were medium range (i.e. within four residues). Only two long-range NOEs (between the $d$ protons of Tyr$\beta$30 and the $b$ and $g$ protons of Gln$\beta$23) were observed, the remainder of the NOEs being intra-residue, as expected for a largely helical peptide. NOE patterns and NOEs per residue are given in Figure 3 and Figure 2, respectively. Because of the considerable overlap in both wild-type and W$\beta$9F spectra, fewer NOEs were used than would be expected, for instance, no $d_{NN}$ NOE was found between residues $-24$ and $-23$. The reason for this absence is simply that the chemical shifts of these two amide protons are so similar that the NOE crosspeak lies too close to the diagonal to be observable. Overlapped NOEs were not used as restraints. The definite absence of some NOEs in the region $-16$ to $-11$ is consistent with a break in the regular secondary structure in this region, in agreement with the CSI results given above.

NOESY spectra acquired at other temperatures (288 K or 298 K) showed no significant differences, although at 288 K, a very weak NOE between Trp$\beta$6 indole NH group and Ile$\beta$5 $H^s$ was visible.

Hydrogen bonds

Of the residues for which the amide-alpha cross-peaks were sufficiently well resolved in TOCSY spectra to measure the peak volume, 21 of the crosspeak ratios (back exchanged:non-exchanged) were below 0.65 and were included as hydrogen bond restraints (Figure 2). It is significant that the amide exchange data are consistent with the CSI and NOE results in suggesting that the central region of the peptide is not a regular helix. On the basis of structures calculated using NOE restraints only, hydrogen bonds were added as restraints between the amide and carbonyl group oxygen atom of the $i-4$ residue. The hydrogen bond restraint from Ile$\beta$5$HN$ to Leu$\beta$10, which is only just within the amide proton exchange cut-off used, consistently violated in structure calculations and so was not included in the final calculations.

Hydrogen exchange was only measured in the W$\beta$9F polypeptide, but 18 hydrogen bonds were included in the wild-type structure calculations for residues N-terminal to His0 as the identical chemical shifts and pattern of NOEs between the two polypeptides in this region imply structural identity (Neuhaus & Williamson, 1989; Potts & Chazin, 1998).

Three-dimensional structure

The NOEs and hydrogen bonds listed above were used as distance restraints in the calculation of 30 structures for each polypeptide. The 14 structures which had the lowest energy were chosen as a representative set for each polypeptide. The mean XPLOR energy for each structure, along with the RMS deviation (rmsd) from ideal geometry and experimental restraints are given in Table 1. The number of NOE violations, and the NOE component to the final energy of the structures, were somewhat higher than was expected for a rigid $\alpha$ helix, probably as a result of flexibility in the molecule causing the average distance between protons to be somewhat greater than would be the case in a rigid molecule. The coordinates of the six lowest energy structures and the XPLOR restraint tables have been deposited in the Protein Data Bank (code 1DX7).
Description of the structure

LH1\(\beta\) is basically helical over most of the molecule; however, the backbone rmsd to the average structure for the whole molecule is high, and the angular order parameters for the \(\phi\) and \(\psi\) backbone angles are low for a central portion of the molecule between residues \(-16\) and \(-8\) (Figure 2). This high rmsd, and the central region of poorly defined backbone angles, implies that the molecule is undefined around the central portion. This is clearly shown in Figure 4, in which the N-terminal and C-terminal portions of the ensemble of structures are separately superimposed.

The polypeptide therefore forms two separate helices (approximately \(-34\) to \(-15\) and \(-11\) to \(+6\)), termed helix 1 and helix 2 respectively, joined by a short undefined region. No difference in crosspeak intensity or line-shape between this more flexible region and the helices was observed in any of the 2D spectra. Most significantly, COSY crosspeak intensities for the central region are very similar to those in the helices. COSY intensities increase in more extended structure and in more mobile regions. The fact that COSY intensities in this region are small, with no change in line-shape, therefore implies that this region does not form an extended or highly mobile structure, but is most probably helical on average.

Ramachandran plots for each polypeptide ensemble show that 91% of backbone angles lie within the favoured regions. Those residues which lie outside this area are at the termini or in the central flexible region. Some backbone angles fall in

Table 1. Structure statistics of LH1\(\beta\) peptides

<table>
<thead>
<tr>
<th></th>
<th>Wild-type</th>
<th>W+9F</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE/res(^a)</td>
<td>9.6</td>
<td>10.4</td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>NOE violations &gt;0.5 Å</td>
<td>2.5 ± 3.0</td>
<td>4.1 ± 3.3</td>
</tr>
<tr>
<td>rmsd from idealised geometry:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.008 ± 0.001</td>
<td>0.009 ± 0.001</td>
</tr>
<tr>
<td>Angle (deg.)</td>
<td>0.845 ± 0.131</td>
<td>0.997 ± 0.168</td>
</tr>
<tr>
<td>Improper (deg.)</td>
<td>0.646 ± 0.134</td>
<td>0.744 ± 0.168</td>
</tr>
<tr>
<td>rmsd from experimental NOEs</td>
<td>0.112 ± 0.015</td>
<td>0.127 ± 0.020</td>
</tr>
<tr>
<td>XPLOR energies(^b) (kcal mol(^{-1}))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>644 ± 194</td>
<td>867 ± 271</td>
</tr>
<tr>
<td>Bonds</td>
<td>49.9 ± 14.2</td>
<td>69.2 ± 19.2</td>
</tr>
<tr>
<td>Angles</td>
<td>154 ± 47</td>
<td>213 ± 69</td>
</tr>
<tr>
<td>Impropers</td>
<td>28.9 ± 11.8</td>
<td>38.0 ± 15.9</td>
</tr>
<tr>
<td>van der Waals</td>
<td>104 ± 57</td>
<td>126 ± 52</td>
</tr>
<tr>
<td>NOEs(^c)</td>
<td>308 ± 85</td>
<td>421 ± 128</td>
</tr>
<tr>
<td>Backbone(^d) rmsd from average (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole molecule ((-34 \rightarrow +6))</td>
<td>3.38</td>
<td>3.02</td>
</tr>
<tr>
<td>Helix 1 ((-33 \rightarrow -19))</td>
<td>0.64</td>
<td>0.83</td>
</tr>
<tr>
<td>Helix 2 ((-11 \rightarrow +5))</td>
<td>0.74</td>
<td>0.70</td>
</tr>
</tbody>
</table>

\(^a\) For residues \(-36\) to \(+10\) (wild-type) and \(-34\) to \(+10\) (W+9F).

\(^b\) Energies were calculated using the CHARMM22 forcefield (Brooks et al., 1983).

\(^c\) An NOE force constant of 50 kcal mol\(^{-1}\) Å\(^{-1}\) was used.

\(^d\) N, C\(^\alpha\) and C\(^\omega\) atoms.
the β-strand region of φ/ψ space, however these angles are those of Asp-36, Pro-8 and Arg+7, also residues at the termini.

The angular order parameter for the χ¹ side-chain angle has a mean value of 0.73 for the wild-type polypeptide (residues -32 to +6, excluding Ala and Gly). These values indicate that the side-chains are poorly defined. The proximity of side-chain chemical shifts observed here to random coil values (Wüthrich, 1986) indicates that the side-chains are conformationally mobile. As the structure is that of an isolated helix, there is no conformational restraint on the side-chains, for instance the side-chain packing found in a hydrophobic core (Creighton, 1993) is absent. The low angular order parameter values for χ¹ are therefore reasonable.

Trp-9 to Phe mutant

This polypeptide, which is unable to form a hydrogen bond to the C3 acetyl (IUPAC nomenclature) of the Bchl (Sturgis et al., 1997) was also assigned and its structure was determined. As expected, the chemical shifts were very similar to those of the wild-type, differing only C-terminal to Ala+2. This high chemical shift identity is indicative of a very similar structure (Potts & Chazin, 1998). Further evidence for a similar structure comes from the very similar pattern of NOEs for both wild-type and W+9F polypeptides. Structure calculations were carried out for the W+9F polypeptide in the same manner as for the wild-type and yielded an essentially identical structure (Figure 4).

Discussion

Validity of the absence of regular helical structure from -14 to -10

The structures calculated reveal a predominantly helical fold to both polypeptides with a more flexible central hinge. There is compelling evidence that this more flexible region is real and not merely an artifact generated by lack of visible NOEs due to overlap. Firstly, the Hα chemical shift index (CSI) indicates a non-helical segment at approximately this position. The CSI-predicted hinge is slightly offset from where the hinge is in the structures, but it must be borne in mind that CSI is not able precisely to predict the origin and termination of secondary structure elements. Secondly, the amide exchange data show a central region in which amide protons are much less protected from exchange with the solvent relative to those protons in clearly helical structure. Thirdly, although there are some NOEs characteristic of helical structure in this region, others are clearly absent (Figure 3), indicating that the regular helical structure is interrupted. However, the line-shape in COSY and TOCSY spectra in this region is the same as for the majority of the protein. This observation, together with the presence of some characteristic helical NOEs, suggests that this region is likely to be predominantly helical, but to flex more than the rest of the protein.

A structure prediction based on sequence alone (Olsen & Hunter, 1994) suggested that there should be a bend in the helix at this point. The divide into two helices has also been predicted (Donnelly & Cogdell, 1993) by studying the periodicity of hydrophilic and hydrophobic residues along the length of a large number of aligned LH polypeptides. Two distinct helical regions were predicted, the first (residues -20 to -9) was suggestive of a water-soluble helix, whilst the second (residues -8 to +6) had the characteristics of a transmembrane helix.

In summary, the more flexible region observed in the LH1β structure is almost certainly a real feature of this molecule in organic solvents. In the membrane and in complex with the other components of the light-harvesting complex the two helices are probably also bent but constrained to a specific conformation with respect to each other and the membrane. It is unlikely that within the membrane these helices fuse to form one continuous helix similar to that seen in LH2 structures.
Organic solvents are known to stabilise the helical conformation and even in some cases induce it (Hirota et al., 1998). Therefore the absence of a well defined helix in this milieu implies that it is unlikely to be formed in vivo.

**Comparison of wild-type and W+9F polypeptides**

A backbone superimposition of the lowest energy structure from each polypeptide ensemble (Figure 4) clearly reveals that both polypeptides share an identical fold and that the mutation does not significantly affect the three-dimensional structures presented. As the point of the mutation is at the penultimate residue, this is not a surprising result. However, the result is of interest because this mutant results in a reduced tendency to form $\beta\beta$ dimers on reconstitution of the LH1 complex rather than $\alpha\beta$ dimers (Kehoe et al., 1998). This change in reconstitution behaviour thus cannot be explained by an intrinsic difference in the structures of the apoproteins. It has been shown (Sturgis et al., 1997) that $\beta$Trp+9 forms a hydrogen bond to Bchl. Therefore it is likely that the change in reconstitution behaviour is due to the loss of the hydrogen bond to the tryptophan, illustrating the importance of intermolecular hydrogen bonding in stabilising LH structures.

**Helix 2 is a transmembrane helix**

The C-terminal helix (residues $-11$ to $+6$) is almost exclusively hydrophobic, containing only four hydrophilic residues, Ser$-7$, His$0$, Tyr$+4$ and Arg$+7$. All of these lie on the Bchl binding face of the helix, and all except the serine residue are known to be involved in pigment:protein interactions (Kehoe et al., 1998; Olsen et al., 1997; Davis et al., 1997). The N-terminal helix is however very different, as it contains a large number of polar and charged residues (see Figure 5). This suggests that helix 2 lies in a more hydrophobic environment than helix 1, as predicted by Donnelly & Cogdell (1993).

Helix 2 is approximately 32 Å in length, and contains five or six turns. This is the same length as the membrane anchoring helix of *Rb. sphaeroides* RC H subunit (Ermler et al., 1994; PDB entry 1PCR), and the transmembrane helices in the RC $L$ and M subunits are typically six turns long. The helix is flanked by tryptophan residues at positions $-10$ and $+6$ (although only Trp$+6$ is conserved in other species), moieties known to be commonly found at membrane/aqueous interfaces (Schiffer et al., 1992). Because Trp$+9$ is known to form a hydrogen bond with Bchl (Sturgis et al., 1997), the $\beta$ polypeptide cannot protrude significantly from the membrane at the C terminus, and therefore this helix is both sufficiently long, and ideally positioned to span the hydrophobic core of the membrane.

**Comparison with Rhodospirillum molischianum LH2β**

The C-terminal region of *Rhodospirillum molischianum* LH2β is highly homologous in
sequence to that of *Rb. sphaeroides* LH1β (Germeroth *et al*., 1993), having 53% sequence identity. Furthermore, its B850 binding site is similar to the *Rb. sphaeroides* B875 binding site (Sturgis *et al*., 1997; Olsen *et al*., 1997; Germeroth *et al*., 1993; Visschers *et al*., 1995). Therefore, one might expect the three-dimensional structure of these polypeptides to be similar, at least around the B875/B850 binding sites.

The C-terminal portions of both polypeptides are very similar. In both cases the helix ends at residue +6 and the C-terminal few residues are non-helical. The C terminus folds back slightly in the *Rs. molischianum* structure into the membrane to ligate the Bchl via Trp+9, but this region in the isolated *Rb. sphaeroides* polypeptide in solution in the absence of the Bchl co-factor is highly mobile, as evidenced by the low angular order parameters (Figure 3). The curvature of the transmembrane helix observed in the NMR ensemble is not present in the *Rs. molischianum* crystal structure and it is probable that interaction with the other components of the complex causes this helix to straighten.

Towards the N terminus, the two structures begin to diverge around residue −11, the start of helix 2 in the *Rb. sphaeroides* structure. The *Rs. molischianum* LH2β polypeptide contains only one long helix which, at the cytoplasmic face, forms the outer edge of the B800 binding site and then rises above the N terminus of the α polypeptide, protruding out of the membrane before terminating at Glu−25. The N terminus lies alongside that of both adjacent α polypeptides, forming hydrogen bonds to them (Koepeke *et al*., 1996). This single long helix which appears to protrude from the membrane is also found in the crystal structure of *Rps. acidophila* LH2 (McDermott *et al*., 1995). If *Rb. sphaeroides* LH1β also formed a single long helix in the same manner, it would extend out from the membrane nine residues further than the helix of *Rs. molischianum*, leaving the N-terminal residues some distance outside the membrane surface.

Thus, the question that arises is whether LH1β in *in vivo* is bent, as in the NMR structure reported here in organic solvents, or straight, as in LH2β. The available evidence points strongly towards a bent structure, as detailed below.

There is little direct evidence on the structure of LH1β in micelles or in *in vivo*. There is a low-resolution projection map of two-dimensional crystals of *Rhodospirillum rubrum* LH1 (Karrasch *et al*., 1995), which is discussed below. The only other structural study is a spectroscopic study of *Rb. sphaeroides* LH1β in detergent micelles (Kikuchi *et al*., 1999), which concluded that both the helical content and the C-terminal structure were closely similar in detergent micelles and in organic solvents. A similar conclusion was reached by Pervushin *et al*. (1994) on bacterioopsin(1-71): they reported that both the secondary and the tertiary structures are similar in organic solvents and in micelles.

There is much indirect evidence that the N terminus of LH1β in *in vivo* must be bent in towards LH1α and the reaction centre. Proteolytic studies show that much of the N terminus is protected from digestion, inconsistent with a projecting helix (Tadros *et al*., 1987). These studies also showed that digestion of LH1β occurs nearer to the N terminus than for LH2β, again implying that the N terminus of LH1β is more shielded from solution by the membrane than is the N terminus of LH2β. Reconstitution studies using truncated polypeptides show that N-terminal truncation of LH1β (i.e. removing residues from the N terminus up to approximately His−18) removes the ability to form intact LH1 complexes, and concluded that the N termini of LH1α and LH1β must interact (Meadows *et al*., 1995, 1998). Geometrical considerations show that the two polypeptides can only interact at the N terminus if the LH1β helix is bent in the centre. Interaction is further indicated as complementarity of charge between the N termini of LH1α and LH1β is required for assembly of the LH1 complex (Loach *et al*., 1994; Dörge *et al*., 1990). Chemical crosslinking studies show that both LH1α and LH1β peptides are close enough to the reaction centre to crosslink to it (Peters *et al*., 1983); only if LH1β is bent can it come close enough to the reaction centre to crosslink. Database and modelling studies have consistently indicated that LH1β is bent in the region of Gly−12 (Olsen & Hunter, 1994; Donnelly & Cogdell, 1993).

Hu & Schulten (1998) proposed a model for the LH1 complex (Hu *et al*., 1998) which is based on the LH2 crystal structure (McDermott *et al*., 1995), but modified to agree with the number of polypeptide subunits seen in LH1 projection maps (Walz *et al*., 1998; Karrasch *et al*., 1995) and is consistent with experimental evidence on the complex obtained from mutagenesis, reconstitution and spectroscopic methods (Sturgis *et al*., 1997; Kehoe *et al*., 1998; Olsen *et al*., 1997, 1994; Davis *et al*., 1997; Meadows *et al*., 1995, 1998). However, this model does not correspond well with the electron density seen in the projection map (Karrasch *et al*., 1995), having relatively too much density in the outer ring, i.e. LH1β (Figure 6). We therefore propose a new model (Figure 7), which is based on the Hu & Schulten (1998) model, with their permission, and differs from it in that the β polypeptide is made up of two distinct helices, of which the N-terminal helix lies at an angle to the membrane normal pointing towards the centre of the ring. This model gives a much better agreement with the experimental projection map (Figure 6), and also allows for greater interaction between the N terminus of LH1β and both LH1α and the reaction centre H subunit. Interestingly, the two-helix structure of LH1β could only be accommodated within the model if the N terminus of LH1α was moved in a...
direction parallel to the membrane plane, suggesting that this region of the \( \alpha \)-polypeptide is also different from that of LH2.

We therefore conclude that the \( \beta \) polypeptide forms a single long \( \alpha \)-helix in LH2, but two helices in LH1. An obvious rationale for this difference is that the \( \beta \) polypeptide is conformationally constrained in the LH2 complex by the presence of the B800 Bchl (which provides a steric block to prevent the polypeptide bending towards the centre of the ring), and also by the requirement that it must interact with LH2\( \alpha \) (which in turn is highly constrained by the hydrogen bond to the magnesium of the B800 Bchl). Neither of these constraints is present in LH1, therefore enabling the N-terminal helix of LH1\( \beta \) to bend at an angle to helix 2 and occupy the space taken up by the B800 Bchl in LH2. The different conformation of the \( \beta \) polypeptides between LH1 and LH2 may also influence the shape of the \( \alpha \beta \) subunit, and hence the overall number of \( \alpha \beta \) subunits in the LH ring. Thus the conformational difference between LH1\( \beta \) and
LH2β presents a mechanism by which the presence of the B800 binding site could simultaneously determine the type of LH complex and the final size of the LH ring.

Materials and Methods

*Rb. sphaeroides* strain DD13, in which the genomic puf and puc operons are absent and the pufBA genes inserted on a plasmid (Jones *et al.*, 1992), was grown on M22+ medium (Hunter & Turner, 1988) and harvested in late log phase. Polypeptides were isolated using the protocol described by Davis *et al.* (1995) and quantified using an extinction coefficient of 3400 M⁻¹ cm⁻¹ per tryptophan residue (Loach *et al.*, 1994). The W+9F mutant strain was a gift from Dr J. Olsen.

Samples were solubilised in CDCl₃:CD₃OH 1:1 (v/v) for NMR by first dissolving in a minimal volume (200-300 μl) of neat trifluoroacetic acid (TFA), which was immediately evaporated under a stream of argon. Mass spectra of TFA-treated samples confirmed that no degradation or covalent modification had occurred (data not shown). The residue was then dissolved in CDCl₃:CD₃OH to a final concentration of approximately 2 mM.

COSY, NOESY and TOCSY spectra were acquired using the States-TPPI method (Marion *et al.*, 1989). All spectra were acquired using a Bruker AMX-500 or DRX-500 MHz spectrometer at 293 K unless otherwise indicated. Data were processed using FELIX 97.0 (Molecular Simulations Inc.). TOCSY spectra (Rance, 1987) were run using SCUBA (Brown *et al.*, 1988) and a mixing time of 80 ms. A total of 512 complex increments were acquired. NOESY spectra (Kumar *et al.*, 1980) were acquired over 1024 complex increments. A mixing time of 200 ms was employed with a composite 180° pulse in the centre of that time to reduce the intensity of artifacts arising from zero-quantum coherence (Neuhaus & Williamson, 1989). DQF-COSY spectra (Rance *et al.*, 1983) were acquired over 650 complex increments. A spectral width of 25 ppm was used throughout.

Distance restraints

Distance restraints were derived manually from NOESY spectra. Even before any structure calculations were performed, the presence of \( d_{\alpha\alpha}(i, i + 3) \) and \( d_{\alpha\beta}(i, i + 1) \) NOEs indicated a helical conformation. Furthermore, the presence of \( d_{\alpha\alpha}(i, i + 4) \) NOEs along with the absence of \( d_{\alpha\alpha}(i, i + 2) \) NOEs indicated that an \( \alpha \) helix, rather than a \( 3_{10} \) helix was present. Upper bound restraints were therefore calculated using a distance calibration obtained by separately averaging the number of contours from several non-overlapped \( d_{\alpha\alpha}, d_{\alpha\beta}(i, i + 1) \) and \( d_{\alpha\gamma}(i, i + 3) \) crosspeaks in regions with a clearly alpha helical pattern of NOEs and comparing to interatomic distances for an \( \alpha \) helix (Neuhaus & Williamson, 1989). Hydrogen bonds were included in structure calculations as distance restraints of \( d(\text{H—O}) \approx 0.5 \) to 2.5 Å and \( d(\text{N—O}) \approx 0.8 \) to 3.5 Å.

Determination of hydrogen bonds

TOCSY spectra were acquired at 288 K on W+9F samples of polypeptide which had either been prepared as above, or had been prepared by first dissolving in TFA-d before dissolving in chloroform:methanol and allowing exchange to occur from ND to NH over nine hours at 288 K. These spectra were normalised relative to a non-exchangeable crosspeak and the ratio of the

**Figure 7.** Model of LH1 based upon the earlier model of Hu *et al.* (1998) and Hu & Schulten (1998) with the authors' permission. The important new feature of this model is that the N-terminal helix of LH1β is directed towards the RC and is predicted to make contact with the RC-H subunit, which occupies the centre of the ring. Alpha and beta polypeptide chains are yellow and magenta, respectively. Bchl molecules are shown in green with the central magnesium atom in red. Three views are shown, from top to bottom; parallel to the membrane plane (cytoplasmic face uppermost), tilted towards the viewer and perpendicular to the membrane plane, viewing from the cytoplasm, respectively.
volumes of amide-alpha crosspeaks between the two spectra measured. Ratios which were less than 0.65 were deemed to be involved in hydrogen bonding, and hydrogen bonds were included in structure calculations for these residues based on structures calculated without these restraints.

Structure calculation

Structures were calculated using XPLOR (Brünger, 1992) to perform a simulated annealing protocol in the manner described (Sorimachi et al., 1996). Some 30 starting structures were created with random coordinates, followed by a simulated annealing protocol. A square-well potential was used for distance restraints and sum averaging employed for equivalent and non-stereo assigned protons. A time step of 1 fs was used throughout. Simulated annealing began using only bond, van der Waals and distance restraint terms over 100 steps of Powell minimisation; 3 ps of dynamics were then performed at 2000 K using chirality and planarity terms in addition to those detailed above. During this time, force constants were increased gradually, with the exception of the repulsive term, which was decreased. The system was cooled in 50 K steps to 100 K during which period of the repulsive term, which was decreased. The system addition to those detailed above. During this time, force

Acknowledgements

This work was supported by funds from the BBSRC and Human Frontier Sciences Program. M.P.W. and M.J.C. wish to thank Molecular Simulations Inc. for the generous provision of FELIX and INSIGHT. We also thank Dr J. Olsen for provision of the W+9F strain, Dr P. Artymiuk for assistance in model building, Dr P. Bulough for help in constructing the projection maps, and Dr K. Schulten for supplying his coordinates of the LH1 model. The Krebs Institute is a Centre for molecular recognition of the Biomolecular Sciences panel of the BBSRC and EPSRC.

References

Donnelly, D. & Cogdell, R. J. (1993). Predicting the point at which transmembrane helices protrude from the bilayer: a model of the antenna complexes from photosynthetic bacteria. Protein Eng., 6, 629-635.
Hunter, C. N. & Turner, G. (1988). Transfer of genes coding for apoproteins of reaction centre and light-


Edited by K. Nagai

(Received 25 November 1999; received in revised form 22 February 2000; accepted 28 February 2000)