



Universiteit
Leiden
The Netherlands

Membrane Protein Studies with Magic Angle Spinning NMR

Gammeren, Adriaan van

Citation

Gammeren, A. van. (2005, April 21). *Membrane Protein Studies with Magic Angle Spinning NMR*. Retrieved from <https://hdl.handle.net/1887/829>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/829>

Note: To cite this publication please use the final published version (if applicable).

Membrane Protein Studies with Magic Angle Spinning NMR

**Sequence specific assignments and structure-function relations
of the photosynthetic light-harvesting 2 complex**

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D. D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 21 april 2005
klokke 16.15 uur

door

Adrianus Josinus van Gammeren

geboren op 29 december 1977

in Waalwijk

Promotiecommissie

Promotor: Prof. dr. H. J. M. de Groot

Co-promotor: Dr. F. B. Hulsbergen

Referent: Prof. dr. H. Oschkinat, FMP Institut, Berlijn

Overige leden: Prof. dr. J. Brouwer
Prof. dr. J. Reedijk
Prof. dr. T. J. Aartsma

Voor Mirjam en mijn ouders

Contents

Preface	7
----------------	----------

Chapter 1

Introduction	9
---------------------	----------

1.1 Solid-state NMR and membrane proteins	9
1.2 Labeling strategies	12
1.3 The light-harvesting 2 complex as model for transmembrane proteins	13
1.4 The BChl <i>a</i> cofactors in LH2	14
1.5 Chemical model systems for the BChl <i>a</i> -histidine complex	15
1.6 Outline of the thesis	16
References	18

Chapter 2

Selective ^{13}C-pattern labeling of the LH2 protein complex	21
---	-----------

2.1 Introduction	21
2.2 Materials and Methods	22
2.3 Results and Discussion	25
2.3.1 Protein labeling pattern	25
2.3.2 BChl <i>a</i> cofactor isotope labeling pattern	31
2.4 Conclusion	31
References	32

Chapter 3

Residual backbone and side-chain ^{13}C and ^{15}N resonance assignments of the LH2 complex	35
--	-----------

3.1 Introduction	35
------------------	----

3.2 Materials and Methods	37
3.3 Results and Discussion	38
3.4 Conclusion	52
References	57

Chapter 4

Resolving the electronic structures of the B800 and B850 cofactors in the LH2 complex 59

4.1 Introduction	59
4.2 Materials and Methods	61
4.3 Results	63
4.4 Discussion	67
4.5 Conclusion	72
References	73

Chapter 5

Chemical model systems for the BChl α -histidine complex 77

5.1 Introduction	77
5.2 Experimental	79
5.2.1 Synthesis	79
5.2.2 (B)Chl α isolation	80
5.2.3 Sample preparations	81
5.2.4 NMR and IR experiments	82
5.3 Results	83
5.3.1 ^1H -NMR	83
5.3.2 ^{13}C -NMR and ^{15}N -NMR	86
5.3.3 IR spectra	90
5.4 Discussion	91
5.5 Conclusion	95
References	95

Chapter 6

General discussion and future outlook	97
6.1 Sequence specific assignments of α -helical transmembrane proteins	97
6.2 Computational methods	99
6.3 Electronic structure of the BChl <i>a</i> cofactors	101
References	103
Summary	105
Samenvatting	109
List of abbreviations	114
Publications	116
Curriculum vitae	117
Nawoord	118

Preface

Proteins are responsible for most reactions in the biochemical pathways of living organisms. Each protein has its own designated function. For example, the protein collagen has a structural role, insulin is a hormonal protein, hemoglobin is used for transport of blood gases and enzymes are involved in food digestion. The analysis of proteins is also of fundamental importance for understanding metabolic pathways and biological functions, including many diseases connected with these pathways.

In this work, Nuclear Magnetic Resonance (NMR) spectroscopy was used, aiming at the sequence specific assignment of the LH2 transmembrane light harvesting 2 (LH2) complex and the characterization of the electronic structure of the embedded cofactors to resolve functional mechanisms. NMR is an analytical technique that can be applied to proteins in liquid or solid state. It allows the detection of chemical shifts from which the spatial high-resolution 3D protein structure can be revealed by extensive data analysis in combination with computational methods, and it is able to determine the electronic structure of cofactors and polar groups. Solution NMR is widely used for water-soluble proteins since the mid-nineteenseventies, while solid state NMR is a relatively young and exclusive technique, which is still in development and can be applied on proteins that are insoluble in water or do not form high quality crystals for X-ray studies. Membrane proteins belong to this class. With almost every life process proceeding sooner or later through a membrane protein, the importance of developing methods to investigate this notoriously difficult class of proteins is difficult to overestimate.

The LH2 transmembrane protein complex with its typical α -helical secondary structure elements and cofactors has been chosen as a representative model, featuring

the characteristics of the relatively unexplored class of transmembrane protein complexes. The LH2 complex is an example of a natural transmembrane protein. A major hurdle in detecting the structure of a membrane protein is to arrive at a sequence specific assignment of the transmembrane helices. As a consequence of the narrow chemical shift dispersion of the backbone nuclei in α -helical transmembrane domains, it is difficult to resolve these shifts in the crowded correlation spectra. The development of novel biosynthetic labeled samples in this work has provided a solution for resolving the chemical shifts of the LH2 complex by solid state NMR spectroscopy. This method can be applied on other membrane proteins as well.

From the assignment of the NMR responses of the bacteriochlorophyll cofactors of the LH2 complex, it is possible to explain their physical properties, like the ground state energy and to detect interactions with the protein matrix. In this way, the NMR study contributes significantly to understanding the mechanistic function of the LH2 complex, which is the light harvesting. This study helps to pave the way for the use of solid state NMR as an independent complementary method to resolve chemical shift data of large α -helical secondary structure domains in transmembrane proteins and to resolve electronic structures. It is anticipated that this method will have many applications on the investigation of these proteins in future.

Chapter 1

Introduction

1.1 Solid-state NMR and membrane proteins

Structure determination of transmembrane proteins by solid state NMR is a relatively novel method and is a technological growth area. Currently, X-ray crystallography and solution NMR spectroscopy are widely used for protein structure determination. They account for more than 25.000 protein structures in the Brookhaven National Laboratories Protein Data Bank.¹ In contrast, only about 50 structures are from membrane proteins. For X-ray studies, difficulties in crystallizing membrane proteins in a lipid environment presents a major obstacle. For liquid state NMR, the long rotational correlation times of proteins incorporated into phospholipid vesicles or bilayers is limiting. Hence, biomolecules that can not be prepared in high-quality crystals or are not soluble in water, like transmembrane proteins, can not be resolved easily by X-ray crystallography and liquid-state NMR spectroscopy, and are relatively unexplored. Because of the development of solid-

state NMR methods and the manufacturing of high-field magnets in the last decade, these molecules can now be studied with solid-state NMR spectroscopy. In this work, the intrinsic transmembrane light-harvesting 2 complex (LH2) from the *Rhodospseudomonas (R.) acidophila* strain 10050 purple bacterium has been chosen as a model protein complex to explore possible routes towards structure determination of membrane proteins by solid-state NMR.

The importance of understanding membrane protein structures and their functions is easy to understand. Membrane proteins often act as enzymes, regulate transport processes, and play a central role in intercellular communication, which makes this class of proteins extremely important. It has been estimated that about 80% of all cellular responses to the external environment are mediated through membrane-bound proteins, including receptors, transporters and channels.² Due to the exponential increase of genomic data in the last decade, statistical sequence analysis has indicated that integral membrane proteins form about 25-30% of all protein sequences.³⁻⁵ Genomic analyses also suggest that relatively small membrane proteins that contain 1 to 4 membrane spanning domains are the most abundant types of membrane proteins. Because of the high abundance of low-molecular weight membrane proteins, solid-state NMR has the potential to become a leading method in the structural proteomics of these structures. Currently, only a few examples of membrane proteins with known crystal structures exist. Most of these are mainly bacterial and plant photoreceptors.⁶⁻⁸ Another example is rhodopsin, the G-protein coupled visual membrane photoreceptor present in the eyes of vertebrates, including the human eye.⁹ However, the difficulty for membrane proteins is that they can not be easily overexpressed and they readily denature, aggregate and need detergents to survive in a functionally active form. Isolation of membrane proteins and conservation by reconstitution into artificial membranes or detergents is not yet routine. For almost every individual membrane protein a specific modified approach is needed to prevent degeneration.²

One of the most critical aspects for understanding the biological functions of proteins is to resolve the interactions between amino acid residues and cofactors, since these are important in mediating a biological process. Many enzymatic chemical processes are not mediated directly by a single residue, but rather by the collective contribution of many functional groups in the active binding site, or

sometimes even by more remote functional sites, which create an optimal chemical or electronic environment for the functionality of the cofactors. Therefore, knowledge about the chemical and electronic protein environment is essential to understand the cofactor function.

With the continued difficulty in crystallizing membrane proteins, solid-state NMR spectroscopy may become an important application in the structure analysis of this important class of proteins. For MAS NMR, local order and a homogeneous environment in the protein sample are sufficient for structure determination, which simplifies the sample preparation procedure. One disadvantage in solid-state NMR spectroscopy is the dominance of anisotropic interactions like chemical shift anisotropy and dipolar couplings. As a consequence of this, the spectral line widths of nuclei in a solid are rather broad, which complicates resolving isotropic chemical shifts of nuclei. The static anisotropic interactions that produce an enormous line-broadening in the solid-state are not present in solution, where these interactions are averaged by the rapid tumbling of molecules.

Solid-state NMR offers two complementary, independent approaches to overcome the anisotropy problem in the solid-state to obtain high-resolution spectra. Both approaches utilize radio frequency (r.f) irradiation to attenuate the many, strong dipolar interactions for line-narrowing and sensitivity enhancement.^{10,11} The first method relies on oriented samples, where peptides or membrane proteins are homogeneously oriented in phospholipid bilayers. The sample orientation provides narrow lines and well-resolved, orientationally dependent spectra. Dipolar couplings and chemical shift interactions depend on the alignment of the molecule relative to the external magnetic field and can therefore be used to study the orientation of the structural elements of the peptide or protein in Magic Angle Oriented Sample Spinning (MAOSS) NMR methods.¹²⁻¹⁵ In addition, it is possible to make distance measurements. The second approach, magic angle spinning (MAS) NMR, is applied to samples containing randomly oriented molecules. In MAS NMR, the chemical shift anisotropy and strong dipolar couplings are significantly reduced when the sample rotates rapidly around the magic angle, $\theta = 54.74^\circ$ with respect to the static magnetic field.¹⁶⁻¹⁸ The line-widths in MAS NMR spectra of ^{15}N and ^{13}C nuclei in protein samples can be reduced below 1.0 ppm. MAS NMR is also applied in this

work for the assignment of the 9×14.2 kDa nonameric light-harvesting 2 complex, which is a first example of a real transmembrane protein complex assignment.

It is generally known that the chemical shift dispersion for ^{13}C and ^{15}N backbone nuclei of α -helical constructed transmembrane proteins is relatively narrow. This yields very crowded correlation spectra, which complicates the extraction of the chemical information and makes sequence specific chemical shift assignment difficult. This contrasts with β -sheet constructed proteins where the chemical shift dispersion is relatively large. To overcome the spectral crowding in the spectra of the α -helical constructed LH2 protein in this work, pattern labeled samples have been developed by using a site-specifically labeled precursor in the expression medium. The preparation of biosynthetically pattern labeled transmembrane protein samples reduces the spectral crowding and facilitates the sequential protein chemical shift assignments. Recent studies have demonstrated that MAS NMR in combination with pattern labeled proteins also opens possibilities to measure long-distance correlations.^{19,20} The pattern labeling method developed here for the LH2 samples also provides long distance correlations, which could not be identified in the crowded spectra of the uniformly labeled LH2 complex. Long-distance correlations are here used predominantly for the sequence specific chemical shift assignment of the long transmembrane α -helical segments in the LH2 complex.

1.2 Labeling strategies

In nature, ^{13}C and ^{15}N isotopes are less abundant. 1,11 % of the carbons and 0,37 % of the nitrogen atoms occur as ^{13}C and ^{15}N isotopes, while the major parts of 98,89 and 99,63 % occur as ^{12}C and ^{14}N isotopes, respectively. Enrichment of to nearly 100 % ^{13}C and ^{15}N isotopes enhances the intensity 90 and 270 fold for carbon and nitrogen NMR signals. The incorporation of isotopes does not change the chemical and physical properties, which offers prospects for the study of larger and more complex biomolecules in their undisturbed native form. In this study, labeling strategies for the LH2 protein, including the labeling of the embedded cofactors, have been developed.

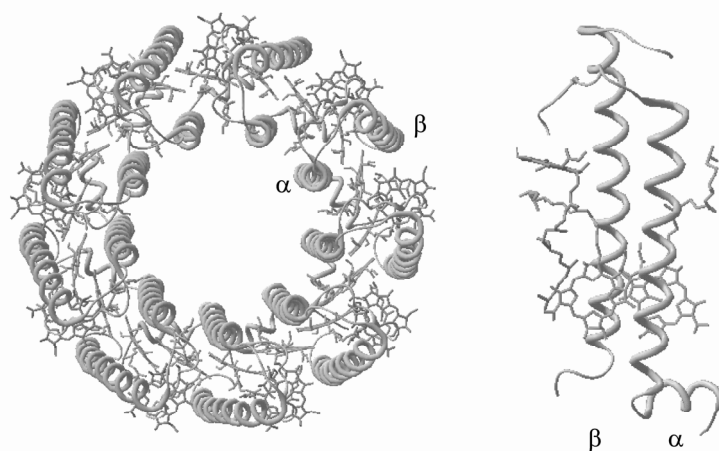
The labeling strategy can be either site-specific or uniform, depending on the target of the study. Specific labeling has mostly been restricted to specific isolated positions of a relatively large protein. A recent example is the specific ^{15}N labeling of the N^π and N^τ atoms in the imidazole ring of the histidines in the LH2 complex to study the coordination of the Mg ion by a histidine residue.²¹ Uniformly labeled protein samples can be obtained by culturing bacteria on a medium containing uniformly ^{13}C and ^{15}N labeled carbon and nitrogen nutrient sources. Extensive selective biosynthetic labeling of a protein, resulting into a labeling pattern as applied to the LH2 complex in this work, should be performed under controlled growth conditions to prevent labeling scrambling. Chapter 2 is devoted to the biosynthesis of pattern labeled protein complexes and the description of the pattern labeling scheme for each residue type and for the BChl *a* cofactors in the LH2 protein.

1.3 The light-harvesting 2 complex as a model for transmembrane proteins

The LH2 complex is an excellent model for studying membrane proteins by MAS NMR spectroscopy, since the LH2 complexes are available in relatively large quantities in the bacterium and can be stabilized with a small amount of detergent after isolation. This enables the implementation of isotope label incorporation starting from a minimum quantity of labeled precursor in the medium. Finally, like many transmembrane proteins, the LH2 protein contains cofactors that are important for mediating the protein function, which can also be studied with solid-state NMR. In addition, a high-resolution (2.0 Å) X-ray structure is known.^{22,23}

The LH2 intrinsic membrane protein complex forms a circular aggregate of nine identical monomeric units with nine-fold symmetry (figure 1.1). Each LH2 monomer consists of two membrane spanning helices, the α -subunit, consisting of 53 residues and the β -subunit consisting of 41 residues. Both subunits comprise a single α -helix spanning the membrane. Each monomer encloses three BChl *a* cofactors denoted by their prominent absorption maxima, B800, αB850 and βB850 ,

and two carotenoids. The α B850 and β B850 form a pair of partly overlapping BChls that is sandwiched between each α - and β -subunit pair in the LH2 monomer. These B850s are axially coordinated at their central Mg ion by β H30 and α H31, respectively. The eighteen B850 cofactors form a continuous overlapping ring in the nonameric structure. The nine B800 cofactors, coordinated at their central Mg ion by the carboxyl- α M1 at the N-terminus of the α -subunit, form a nine-membered ring without overlap.



α : MNQGK⁵IWTV¹⁰NPAIG¹⁵IPALL²⁰GSVTV²⁵IAILV³⁰HLAIL³⁵SHTTW⁴⁰FPAYW⁴⁵QGGVK⁵⁰KAA

β : ATLTA⁵EQSEE¹⁰LHKYV¹⁵IDGTR²⁰VFLGL²⁵ALVAH³⁰FLAFS³⁵ATPWL⁴⁰H

Figure 1.1 Left: The X-ray structure of the nonameric LH2 complex from *Rhodospseudomonas acidophila* strain 10050. Right: the structure of the monomeric LH2 complex derived from the 1NKZ²³ PDB file. Bottom: the primary sequences of the α - and β -subunit of the monomeric LH2 complex; membrane spanning residues are underlined.

1.4 The BChl α cofactors in LH2

Photosynthetic bacteria have developed highly efficient protein complexes to collect sunlight and to use this for their metabolic processes. This energy-conversion process, better known as the bacterial photosynthesis, is accomplished

through the interplay of different protein complexes of the bacterial photosynthetic unit. The interactions between the cofactors and the interactions of the cofactors with the proteins are of major importance. The first step of bacterial photosynthesis is the absorption of a photon by a BChl or carotenoid in the LH2 complex, followed by an ultra fast transfer of the excitation energy between BChls and Förster type energy transfer to the BChl of the bacterial reaction center. For a convenient review see ref.²⁴ The excited BChl in the reaction center is oxidized by the protein complex, yielding a charge separation, which is further stabilized by the protein complex.

A part of this thesis is devoted to the study of the electronic ground states of the BChl *a* cofactors of the of the LH2 complex in relation to their maximum absorption wavelength (Chapter 4). The main absorption maxima of the B800 and the B850 are red-shifted compared to the main absorption maximum of isolated monomeric BChl *a* in acetone solution.²⁵ The MAS NMR study on these cofactors has provided complete chemical shift datasets for each distinct BChl *a* cofactor in its native environment. Density Functional Theory (DFT) calculations have been used to compensate the observed chemical shifts for the ring current effects from the aromatic BChl *a* macrocycles. Subsequently, based on the chemical shift changes, an electronic basis for the different main absorption maxima of the BChl *a* cofactors has been derived, explaining the electronic ground state of the BChl *a* cofactors.

1.5 Chemical model systems for the BChl *a*-histidine complex

An important interaction in the LH2 complex is the axial coordination bond between the imidazole moieties of α H31 and β H30 and the central Mg-ion of the BChl cofactors. It is suspected that this key interaction is not only structurally but also electronically important for the function of the complex. For a detailed insight into the interactions between the protein binding site, the cofactors, and the electronic structure of the histidine-BChl *a* complex, novel chemical models have been developed. Magnesium ions have been incorporated into octa-ethyl porphyrin (OEP) to model the BChl *a* cofactor. In parallel, also BChl *a* cofactors have been isolated from photosynthetic bacteria for modeling studies. The axial coordination of the α H31 and β H30 residue has been mimicked with an imidazole (Im) or 1-methyl

imidazole (1-MeIm). Solution NMR, solid-state MAS NMR and Infra Red (IR) spectroscopy have been used to investigate the stoichiometry of the model complexes and the electronic properties of the axial ligand and the BChl *a* bacteriochlorophyll ring.

1.6 Outline of the thesis

In this thesis the sequence specific assignment of the transmembrane LH2 complex has been performed, which is an essential step to determine the structure of the protein. Also the electronic characterization of the embedded BChls in the ground state has been worked out. MAS NMR spectroscopy, isotropic chemical shift analysis and the development of advanced chemical models have been used to arrive at this point.

For sequence specific assignment purposes, a biosynthetically site-specific pattern labeling approach in combination with MAS NMR spectroscopy has been implemented aiming at resolving residue-specific chemical shifts of α -helical segments in the LH2 complex as initial step. A second aim of this thesis is electronic characterization of the BChl cofactors in LH2 by resolving the interactions with the protein matrix. This aims to achieve a mechanistic understanding of the protein function of light-harvesting. Chemical models, which mimic the function, structure and spectroscopy of the intact LH2 complex, have been designed for a detailed understanding of the structural and electronic role of the coordination bonds that bind the cofactors to their binding site.

Chapter 2 describes the preparation of pattern labeled samples. Various pattern labeled samples have been obtained by adding site-specifically ^{13}C -labeled succinic acid, labeled amino acids and ^{15}N -labeled NH_4OH to the expression medium under controlled growth conditions. The site-specific ^{13}C -labeled nutrient sources have been consumed and incorporated into the LH2 protein following mainly a single metabolic pathway. The labeling patterns of the amino acids in the pattern labeled LH2 protein have been deduced with 2D ^{13}C - ^{13}C and ^{13}C - ^{15}N correlation spectra. It is demonstrated that MAS NMR experiments in combination with the use of pattern

labeled protein samples facilitate the NMR chemical shift assignment of α -helical segments in the LH2 complex.

Various pattern labeled samples are described in chapter 2 and used to perform a sequence specific assignment, which is described in **Chapter 3**. The assignment procedure is based on a collection of 2D MAS NMR spectra containing many ^{13}C - ^{13}C and ^{13}C - ^{15}N intra- and inter-residue correlations. The procedure for the sequential assignment of the LH2 protein has been described in detail and it has been demonstrated how the applied method benefits the investigation of transmembrane proteins with MAS NMR spectroscopy.

Chapter 4 focuses on the electronic ground state of the three BChl cofactors in their natural environment. 2D RFDR and PDS MAS NMR experiments enabled a distinct chemical shift assignment for each of the three BChls, i.e. B800, α -B850 and β -B850. Analyzing the isotropic chemical shifts from the MAS NMR correlation spectra for each carbon involved has revealed the electronic changes of the different types of embedded BChls with respect to their monomeric forms. The NMR data obtained have been used to resolve the interactions between the cofactors and to resolve the interactions of the cofactors with the protein matrix. The results help to explain the red-shift of the BChls in their natural environment relative to their monomeric forms *in vitro*.

Chapter 5 deals with chemical models that probe the spectroscopic features of the coordination bond between the central Mg ions of the BChls and the imidazole moieties of the histidine residues. This important bond contributes to the partial charge transfer from the imidazole moiety to the conjugated macrocycle of the BChl *a* cofactor. Additional insight into the coordination structure and chemical properties has been obtained by NMR and IR spectroscopy. The spectroscopic features of the chemical models can be directly related to the spectroscopic features of the natural LH2 system and are discussed in detail.

Chapter 6 provides a general discussion on the results obtained and considers possible strategies for future experiments and recommendations to resolve the structure of membrane proteins by MAS NMR. Also the importance of a computational aid for MAS NMR studies on transmembrane proteins has been worked out. Finally, the importance of the resolved electronic ground state structures of the BChls and the role of MAS NMR to resolve electronic cofactor

structures have been discussed. Major parts of chapters 2, 3, 4 and 5 have been published.²⁶⁻²⁹

References

- (1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E.; *Nucleic Acids Res.* 2000, 28, 235-242.
- (2) Watts, A.; Burnett, I. J.; Glaubitz, C.; Grobner, G.; Middleton, D. A.; Spooner, P. J. R.; Watts, J. A.; Williamson, P. T. F.; *Nat. Prod. Rep.* 1999, 16, 419-423.
- (3) Simon, I.; Fiser, A.; Tuszny, G. E.; *BBA-Protein Struct. M.* 2001, 1549, 123-136.
- (4) Wallin, E.; von Heijne, G.; *Protein Sci.* 1998, 7, 1029-1038.
- (5) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. L.; *J. Mol. Biol.* 2001, 305, 567-580.
- (6) Koepke, J.; Hu, X. C.; Muenke, C.; Schulten, K.; Michel, H.; *Structure* 1996, 4, 581-597.
- (7) Kuhlbrandt, W.; *Nature* 2001, 411, 896-899.
- (8) Prince, S. M.; Papiz, M. Z.; Freer, A. A.; McDermott, G.; HawthornthwaiteLawless, A. M.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 1997, 268, 412-423.
- (9) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M.; *Science* 2000, 289, 739-745.
- (10) Waugh, J. S.; Huber, L. M.; Haeberle, U.; *Phys. Rev. Lett.* 1968, 20, 180-182.
- (11) Pines, A.; Gibby, M. G.; Waugh, J. S.; *J. Chem. Phys.* 1973, 59, 569-590.
- (12) Opella, S. J.; Marassi, F. M.; *Chem. Rev.* 2004, 104, 3587-3606.
- (13) Opella, S. J.; Nevzorov, A.; Mesleh, M. F.; Marassi, F. M.; *Biochem. Cell Biol.* 2002, 80, 597-604.
- (14) Sizun, C.; Bechinger, B.; *J. Am. Chem. Soc.* 2002, 124, 1146-1147.
- (15) Vosegaard, T.; Nielsen, N. C.; *J. Biomol. NMR* 2002, 22, 225-247.
- (16) Andrew, E. R.; Bradbury, A.; Eades, R. G.; *Nature* 1959, 183, 1802-1803.

-
- (17) Andrew, E. R.; Bradbury, A.; Eades, R. G.; *Nature* 1958, 182, 1659-1659.
- (18) Andrew, E. R.; Newing, R. A.; *P. Phys. Soc. Lond.* 1958, 72, 959-972.
- (19) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H.; *Nature* 2002, 420, 98-102.
- (20) Castellani, F.; van Rossum, B. J.; Diehl, A.; Rehbein, K.; Oschkinat, H.; *Biochemistry* 2003, 42, 11476-11483.
- (21) Soede-Huijbrechts, C.; Cappon, J. J.; Boender, G. J.; Gast, P.; Hoff, A. J.; Lugtenburg, J.; de Groot, H. J. M.; In: *Photosynthesis: mechanisms and effects.*; Kluwer Academic Publ.: Dordrecht, 1998.
- (22) McDermott, G.; Prince, S. M.; Freer, A. A.; Hawthornthwaite-Lawless, A. M.; Papiz, M. Z.; Cogdell, R. J.; Isaacs, N. W.; *Nature* 1995, 374, 517-521.
- (23) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 2003, 326, 1523-1538.
- (24) Hu, X. C.; Ritz, T.; Damjanovic, A.; Autenrieth, F.; Schulten, K.; *Q. Rev. Biophys.* 2002, 35, 1-62.
- (25) Scheer, H.; In: *The Chlorophylls.* CRC Press: Boca Raton FL, 1991.
- (26) van Gammeren, A. J.; Hulsbergen, F. B.; Hollander, J. G. & de Groot, H. J. M., *J. of Biomol. NMR* 2004, 30, 267-274.
- (27) van Gammeren, A. J.; Hulsbergen, F. B.; Hollander, J. G. & de Groot, H. J. M., *Journal of Biomolecular NMR* 2005, *in press*.
- (28) van Gammeren, A. J.; Buda F., Hulsbergen, F. B.; Kiihne S.; Hollander, J. G.; Egorova-Zachernyuk, T. A.; Fraser, N. J.; Cogdell, R. J. & de Groot, H. J. M., *J. Am. Chem. Soc.* 2005, *in press*.
- (29) van Gammeren, A. J.; Hulsbergen, F. B.; Erkelens, C. & de Groot, H. J. M.; *J. Biol. Inorg. Chem.* 2004, 109-117.

Chapter 2

Selective ^{13}C -pattern labeling of the LH2 protein complex*

2.1 Introduction

At an early stage, the light-harvesting 2 (LH2) transmembrane protein complex from the anaerobic *Rhodospseudomonas (R.) acidophila* strain 10050 purple bacterium has been used to explore the range and resolution of the MAS NMR in the study of transmembrane proteins.¹ The LH2 complex, for which a detailed X-ray structure has been described elsewhere,²⁻⁴ and briefly in chapter 1, is employed for selective labeling methods to investigate the protein structure by solid-state NMR. Two-dimensional proton driven spin diffusion solid-state NMR correlation spectroscopy has been used to trace each individual ^{13}C isotope from the labeled succinic acid precursor to its destination into the protein and into the embedded major light-absorbing bacteriochlorophyll cofactors.

MAS NMR applied to biomolecules requires suitable [^{13}C , ^{15}N]-labeled samples of high structural homogeneity for optimal resolution, to resolve the large number of

* The contents of this chapter have been published in the *Journal of Biomolecular NMR*, 30, 2004, 267-274.

cross signals from the backbone nuclei of the protein in dipolar correlation spectra. Transmembrane proteins generally contain α -helices and there is little dispersion for the ^{13}C and ^{15}N resonances of backbone nuclei in the α -helical structure, which makes the sequence specific chemical shifts assignments difficult. A reduction of labeled carbon positions is a prerequisite to resolve the correlation responses in 2D spectra.

This chapter describes the preparation of selectively pattern labeled LH2 samples in an attempt to address the problem of spectral crowding for chemical shift assignments and to suppress the broadening of signals due to J-couplings between neighboring carbons for an increased spectral resolution. The use of site-directed reduced biosynthetic labeled samples has been demonstrated before with *E. coli* and has appeared to be useful for the structure determination of the α -spectrin SH3 domain.^{5,6} The proteins of this system were pattern labeled by using specifically labeled [1,3- ^{13}C] or [2- ^{13}C] glycerol as carbon a source and it has been established how the labels were incorporated into the protein residues.⁶ The anaerobic *R. acidophila* purple bacterium does not adapt to the medium with glycerol as a single carbon source and a different experimental labeling method is needed. For biosynthetic incorporation of ^{13}C nuclei into the LH2 complex, chemically synthesized site-specific labeled [1,4- ^{13}C] and [2,3- ^{13}C]-succinic acid have been used, yielding the 1,4-LH2 and 2,3-LH2 sample, respectively. Due to the symmetry of the succinic acid precursor, using a singly labeled or [1,3- ^{13}C]-succinic acid yields a 50% dilution of the labeled positions in the biomolecules. The labeling patterns of the 1,4-LH2 and 2,3-LH2 and the embedded BChl *a* cofactors have been analyzed using 2D PDSD dipolar correlation NMR spectra and are described in this chapter.

2.2 Materials and methods

All [^{13}C , ^{15}N] isotopically enriched LH2 complexes were obtained by growing the photosynthetic purple bacteria *R. acidophila* strain 10050 anaerobically in light at 30 °C on a well-defined medium.¹ For an optimal cell growth, *R. acidophila* strain

10050 requires an elaborate medium containing two essential nutrient sources, consisting of succinic acid and a mixture of amino acids (Cambridge Isotopes Laboratories, Andover, MA, USA). ^{13}C -labeled succinic acid was prepared by a multi-step synthesis, starting from ^{13}C -labeled acetic acid.⁷ In the final step, fumaric acid was converted into succinic acid using a reduction procedure with H_2 and Pd-C as catalyst.

The label incorporation from the two different nutrient sources was investigated by labeling one or both carbon nutrient sources. Three labeled LH2 samples were prepared: the uniformly ^{13}C -labeled LH2 sample (U-LH2), the 1,2,3,4-LH2 sample that was prepared from a uniformly ^{13}C labeled [1,2,3,4- ^{13}C]-succinic acid and the AA-LH2 sample that was prepared from a [^{13}C , ^{15}N]-labeled amino acid mixture. For uniformly labeled samples the metabolic pathway is irrelevant. The concentrations of the [^{13}C , ^{15}N]-isotope labeled amino acid mixture and [1,2,3,4- ^{13}C]-succinic acid in the medium were optimized to 1.5 g of amino acid mixture and 2.0 g succinic acid per liter to minimize the label cost.

In contrast, for the preparation of specific biosynthetic labeled samples, the selection of one specific pathway to minimize the dilution or scrambling of the label pattern in the protein by synthesis of the same intermediates by different competitive non-equivalent metabolic pathways is required. For the preparation of the 1,2,3,4-LH2 sample, the concentration of the unlabeled amino acid mixture was reduced to 1.0 g/l to enhance the biosynthesis from the labeled succinic acid source and to reduce the incorporation of unlabeled amino acids. For the AA-LH2 sample, both the concentrations of the unlabeled succinic acid and the labeled amino acid mixture were adjusted to 1.5 g/l to enhance the label incorporation from the amino acid source. In a preliminary try-out experiment, an AA-LH2 sample was prepared from the hydrolysate of waste proteins from the *E. Coli BL21*, which was used for the expression of the SH3 protein.⁵ The uptake of residues from the mixture was insufficient to obtain an extensive labeled AA-LH2 sample.

A proton driven spin diffusion (PDSF) spectrum was recorded on these three test samples to gain insight into how nutrient sources are involved in the biosynthesis of residues in the LH2 complex. It was found that succinic acid is required for the synthesis of most residues. Subsequently, the pattern labeled 1,4-LH2 and 2,3-LH2 samples were prepared from a medium containing 2.0 g/l [1,4- ^{13}C]-succinic acid

and [2,3-¹³C]-succinic acid, respectively, to introduce a labeling pattern. The amount of unlabeled amino acid mixture in the media was reduced to 1.0 g/l to enhance the biosynthesis from the labeled succinic acid source and to reduce the incorporation of unlabeled amino acids. All media, including the three test samples, contained also ¹⁵NH₄OH. The concentration was twice the molar succinic acid concentration for labeling all ¹⁵N atoms in the protein complex.

The cell growth was followed with UV spectroscopy by measuring the optical density at $\lambda = 860$ (OD₈₆₀). A full-grown culture will enable alternative biosynthetic pathways to maintain its steady state, which may lead to dilution or scrambling of the labels. To prevent deterioration of the labeling pattern, cells obtained from media containing site-specific labeled succinic acid were harvested at OD₈₆₅ = 3.6/cm, which is just before the cell culture reaches the steady state (OD₈₆₅ \geq 3.8). The LH2 complex isolation was performed according to Hawthornthwaite-Lawless and Cogdell.⁸ About 30 mg of LH2 protein was isolated, starting from 0.8 l medium. The amount of protein was used to fill two 4 mm CRAMPS rotors.

1D and 2D ¹³C-¹³C homonuclear correlation spectra of labeled LH2 samples were recorded using PDS MAS NMR spectroscopy with a Bruker AV-750 spectrometer equipped with a double channel CP-MAS probe head and with a ¹³C radio frequency of 188 MHz. The proton $\pi/2$ pulse was set to 3.1 μ s, corresponding with a nutation frequency of 80.6 kHz. ¹³C B₁ field strengths of 50 kHz corresponding with a cross polarization time of 2.0 ms were used during a 100-50% ramped CP sequence.⁹ In the PDS experiment, two-pulse phase modulation (TPPM) decoupling was applied during the t_1 and t_2 periods.¹⁰ A mixing time of 50 ms was used to transfer the magnetization into the side chains. All samples were cooled to 253 K, and the MAS spin frequency $\omega_R/2\pi$ was 8.5 kHz. The ¹³CO resonance of U-[¹³C,¹⁵N]-Tyrosine·HCl at 172.1 ppm was used as an external reference for the determination of the isotropic ¹³C chemical shifts.

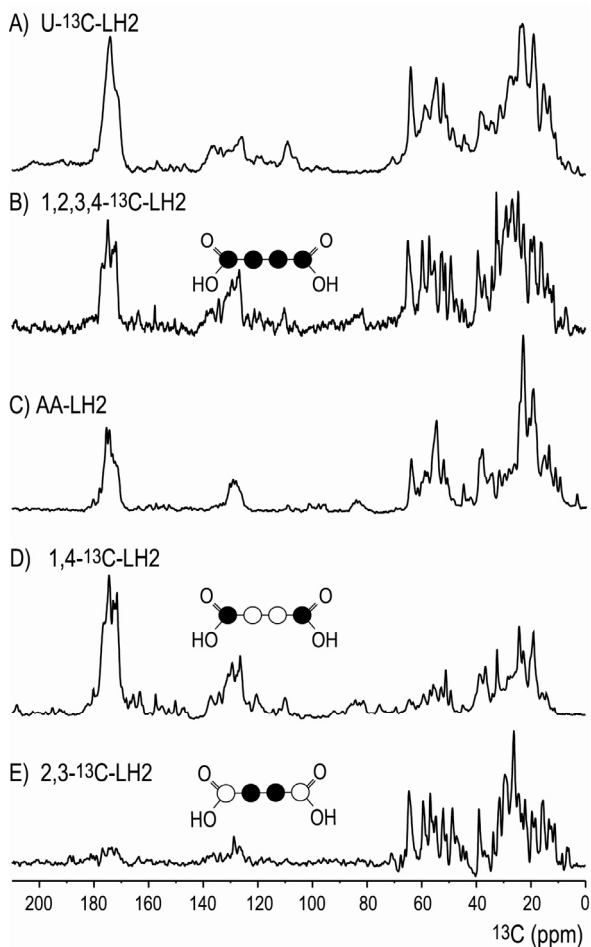


Figure 2.1 1D PDSO spectra of U-LH2 (A), 1,2,3,4-LH2 (B), AA-LH2, 1,4-LH2 (D) and 2,3-LH2 (E) recorded with 32 scans. A schematic representation of [1,4- ^{13}C]-succinic acid (D) and [2,3- ^{13}C]-succinic acid (E) indicate the selective LH2 labeling used to obtain the 1D spectra.

2.3 Results and discussion

2.3.1 Protein labeling pattern

Five 1D ^{13}C PDSO spectra have been recorded from U-LH2, 1,2,3,4-LH2, AA-LH2, 1,4-LH2 and 2,3-LH2 complexes (figure 2.1A-E). Resonances between 0 and 75 ppm are mainly from the labeled $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma$ and $\text{C}\delta$ carbons in the amino acids. A minor

fraction is from aliphatic carbons of the BChl *a* cofactors. Between 100 and 170 ppm the signals from aromatic amino acids and the bacteriochlorin rings of the BChl *a* cofactors are detected. The carbonyl responses from the protein backbone and some carbonyl responses from the BChl *a* cofactors are located between 170 and 190 ppm.

The 1D spectra of U-LH2, 1,2,3,4-LH2 and AA-LH2 complexes in figure 2.1A, 2.1B and 2.1C all show resonances in all three spectral regions, demonstrating that in particular the succinic acid is essential for labeling the protein. The spectrum of the AA-LH2 in figure 2.1C shows both signals in the carbonyl area and in the aliphatic area, indicating that also a part of the labeled amino acids are introduced from the labeled amino acid mixture nutrient source. Resonances of the 1,4-LH2 and the 2,3-LH2 complexes are represented in figures 2.1D and 2.1E, respectively. In the 1,4-LH2 spectrum in figure 2.1D, intense resonances are observed in the carbonyl and aromatic regions, while relatively weak responses are observed in these regions for the 2,3-LH2 sample in figure 2.1E. Contrary, many responses are observed in the aliphatic part of the 2,3-LH2 dataset, while for the 1,4-LH2 a less crowded spectrum is observed in this area.

The 2D PDSO correlation spectra reveal the labeled pairs of nuclei that are in the molecular structure of each sample. From this information it was deduced which residues are synthesized from [1,2,3,4-¹³C]-succinic acid and which residues are taken up from the amino acid mixture, and how the ¹³C labels of [1,2,3,4-¹³C]-succinic acid are incorporated in the BChl *a* cofactors. The aliphatic part of the 2D PDSO spectrum between 0 and 75 ppm shows the cross signals between carbons of the amino acid side chains. A complete and specific set of correlations for almost every residue except I and L has been detected. I and L residues comprise about 20% of the protein sequence and are responsible for a significant reduction of the correlations in the 2D spectra of the samples prepared by succinic acid. In the protein there is one carboxyl-methionine at the N-terminus of the α -subunit, which can not be identified in the spectrum. The data show that almost all amino acids, except I and L, are synthesized from the succinic acid carbon source. This is corroborated by a 2D spectrum of the AA-LH2 sample, which shows the correlation signals for I and L. The same spectrum also shows resonance sets for V, G and A, which can be taken up from the amino acid mixture or

synthesized from the succinic acid source. By using separately both the AA-LH2 and 2,3-LH2 samples, it is possible to observe all type of amino acids, which will be used for the assignments, described in chapter 3.

The aliphatic parts of the 2D PDS spectra of 1,4-LH2 and 2,3-LH2 in figure 2.2 (p. 28) show a reduced number of cross signals compared to data collected from the 1,2,3,4-LH2. From the 1D PDS spectra in figure 2.1D and 2.1E, it was inferred that most of the labels from [1,4- ^{13}C] succinic acid are introduced at the carbonyl positions, while the labels from [2,3- ^{13}C] succinic acid are mainly introduced at the $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma$ and $\text{C}\delta$ positions. This is confirmed by the aliphatic correlation area in the 2D PDS spectra of the 2,3-LH2 and 1,4-LH2 samples in figure 2.2. The $\text{C}\alpha$ - $\text{C}\beta$ correlation area in spectrum of the 2,3-LH2 complex is more crowded than for the 1,4-LH2 complex. This demonstrates that the [2,3- ^{13}C]-succinic acid labels are mainly introduced into the aliphatic $\text{C}\alpha$, $\text{C}\beta$, $\text{C}\gamma$ and $\text{C}\delta$ positions.

The reverse is observed for the CO- $\text{C}\alpha$ - $\text{C}\beta$ correlation area in the spectra of the 1,4-LH2 and 2,3-LH2 in figure 2.3 (p. 29), where the carbonyl response area for the 1,4-LH2 is more crowded than for the 2,3-LH2 sample. This confirms that the [1,4- ^{13}C]-succinic acid labels are mainly incorporated in the carbonyls. The relatively weak cross resonances of the 1,4-LH2 complex indicated in grey in the CO- $\text{C}\alpha$ - $\text{C}\beta$ correlation area between 170-180 ppm in figure 2.3, represented the H, T, S, V, A and P residues in which $\text{C}\alpha$ carbons are partly enriched by label scrambling. The $\text{C}\alpha$ carbons that correlate with the carbonyls are labeled from [1,4- ^{13}C]-succinic acid. The $\text{C}\alpha$ -CO responses are weak, indicating fractional labeling to multiple metabolic pathways. The carbonyls of the histidines are anomalous in the sense that the major fraction is labeled from [2,3- ^{13}C]-succinic acid. This can be concluded from the relatively strong CO- $\text{C}\alpha$ - $\text{C}\beta$ responses in the spectrum of the 2,3-LH2 complex. Weaker CO- $\text{C}\alpha$ cross correlation signals are detected with the same shifts in the spectrum of the 1,4-LH2. This indicates that a fraction of the CO and $\text{C}\alpha$ carbons of the histidines can be enriched from [1,4- ^{13}C]-succinic acid. The $\text{C}\alpha$ - $\text{C}\beta$ correlation signals of the histidine residues are clearly observed in the spectrum of the 1,4-LH2 (figure 2.2). The few remaining cross correlations in the spectrum of the 2,3-LH2 in figure 2.3 are attributed to the labeled CO, the $\text{C}\alpha$, the $\text{C}\beta$ and the $\text{C}\gamma$ carbons of the Q and E residues.

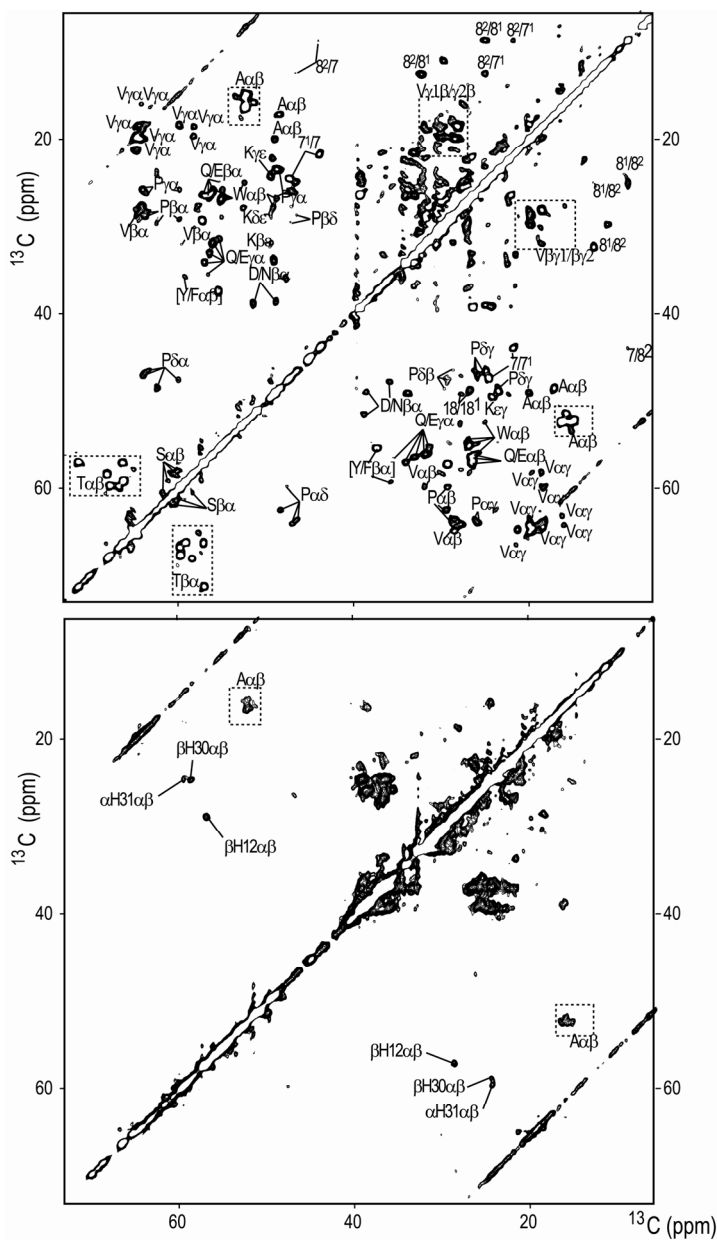


Figure 2.2 The aliphatic parts of the PSD 2D spectra collected from the 2,3-LH2 (top) and the 1,4-LH2 sample (bottom). Correlation signals are indicated by their a single letter code of the corresponding amino acid residue. The resonances of the BChl a cofactors are indicated with a number corresponding to the IUPAC numbering of BChl a as represented in figure 2.4.

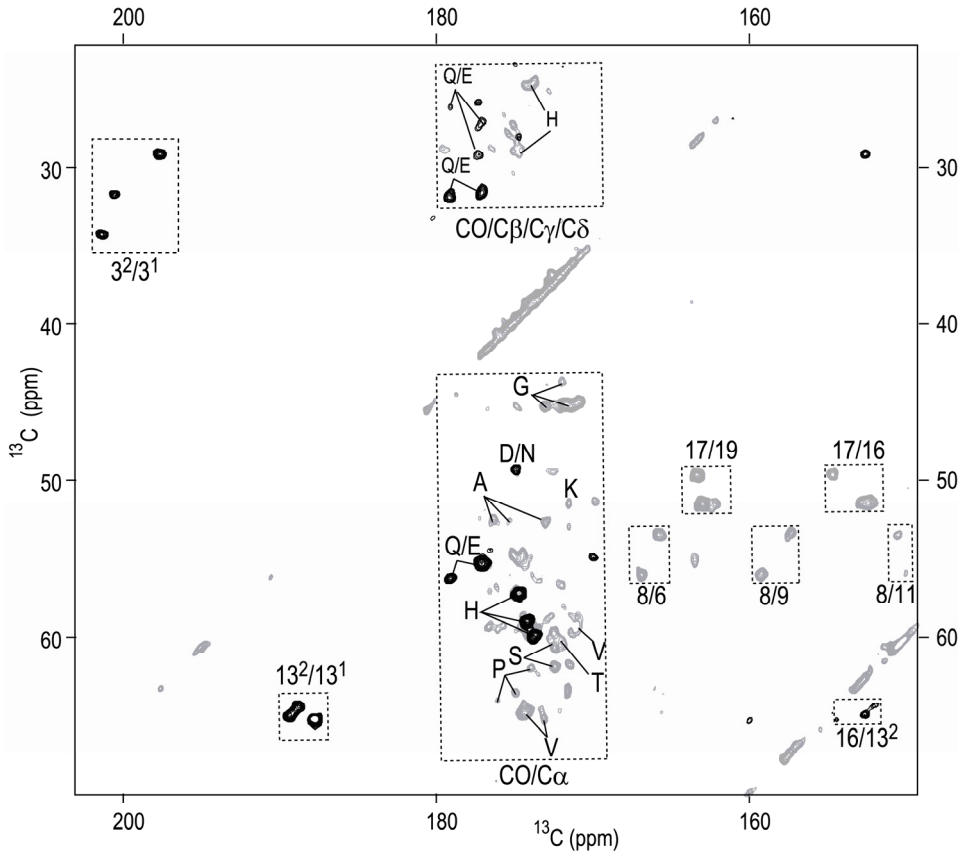


Figure 2.3 The superposition of the CO-C α -C β -(C γ) correlation areas in the data obtained from 1,4-LH2 (grey) and 2,3-LH2 (black). Signals from the BChl *a* cofactors are numbered corresponding to the IUPAC carbon numbering of BChl *a* as represented in figure 2.4.

Finally, by comparing the PDS spectra of the 2,3-LH2 and 1,4-LH2 with the corresponding spectra of the U-LH2 and the 1,2,3,4-LH2, the isotope incorporation for each amino acid can be determined. The label incorporation is summarized in figure 2.4, which represents the label transfer from succinic acid to the ^{13}C positions for each amino acid type by black and grey circles. The small sections in the circles indicate minor fractions of ^{13}C -label rearrangement due to the biosynthetic pathways that are different from the main pathway. As indicated by the relative site-specific biosynthetic incorporation in the figure, the succinic acid digestion follows mainly a single metabolic pathway. Evidently, the succinic acid

concentration of 2.0 g/l medium to prepare the 1,2,3,4-LH2, the 1,4-LH2 and the 2,3-LH2 sample has appeared to be sufficiently high to suppress the alternative metabolic pathways.

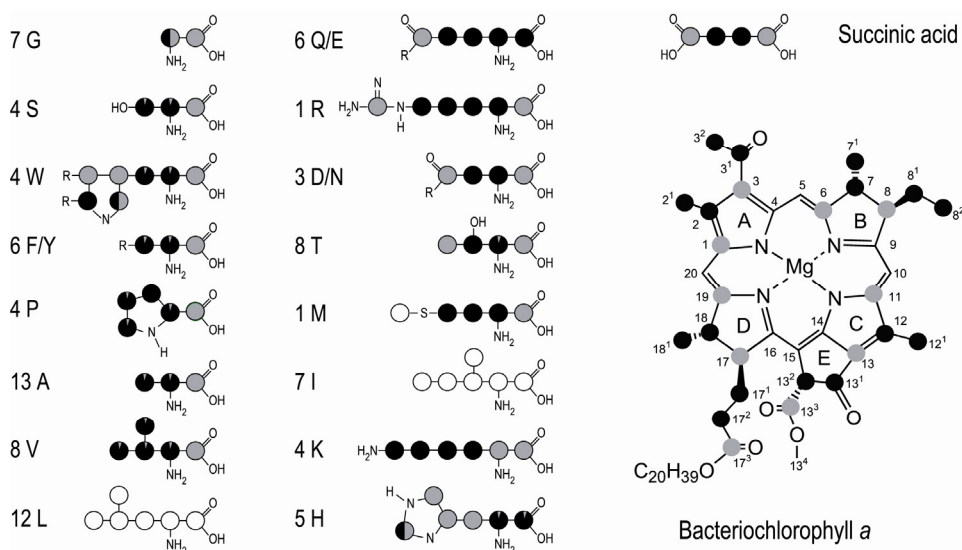


Figure 2.4 Schematic representation of the effective ^{13}C -isotope enrichment of the residues and the BChl a cofactors, obtained by protein expression in *R. acidophila* strain 10050. The grey circles correspond to the ^{13}C labeling pattern that is obtained by growing on $[1,4\text{-}^{13}\text{C}]$ -succinic acid, while the labeling pattern obtained by growing on $[2,3\text{-}^{13}\text{C}]$ -succinic acid is represented with black circles. The small section indicates a small fraction of label scrambling.

The resonances in the pattern labeled 2,3-LH2 and 1,4-LH2 has been significantly reduced. It has been shown for a pattern labeled β -sheet protein that the decreased number of correlation signals in the pattern labeled sample in combination with MAS NMR techniques offers a simplified route to a sequence specific assignment of membrane proteins by MAS NMR.^{5,11} The pattern labeling in the 1,4-LH2 and 2,3-LH2 complexes can also be used for a sequence specific assignment of the transmembrane LH2 complex, which will be described in chapter 3.

2.3.2 BChl *a* cofactor isotope labeling pattern

From figure 2.1D and 2.1E, it transpires that most aromatic carbons of the BChl *a* cofactors are labeled from the [1,4- ^{13}C]-succinic acid. This is confirmed by the grey aromatic cross correlations in the 2D spectra in figure 2.3. The numbering and color coding for the BChl *a* follows the scheme of figure 2.4. Strong C2-C2¹ and C12-C12¹ are detected in the dataset collected from the 2,3-LH2 outside the response areas represented in figures 2.2 and 2.3. The carbon positions in figure 2.4 that are not represented by circles are introduced by glycine in the biosynthesis of BChl *a*. They are slightly enriched by a small fraction of glycine synthesized from the labeled succinic acid, which is in competition with the main unlabeled glycine pool from the amino acid mixture.

The level of enrichment is generally too low for the detection of a cross signal. Only C9 and C16 are weakly observed in the spectra. The observation of positions C9 and C16 is probably owing to the relatively close distance from the aliphatic groups at C8 and C17, respectively, where the multiple ^1H nuclei at these aliphatic carbons mediate the magnetization transfer in the PDS MAS NMR technique. Positions C4, C5, C10, C14, C15 and C20 are remote from fully labeled hydrogenated carbons, which make an observation of correlation signals unlikely. The labeling pattern of the BChl *a* cofactors in the 1,4-LH2 and 2,3-LH2 complexes as represented in figure 2.4 is well in line with the biosynthetic condensation pathway of succinyl-CoA and glycine to δ -aminoleuvalinic (δ -ALA) acid and from δ -ALA to BChl *a*.¹² A retro-synthesis from BChl *a* to succinic acid can reduce the labeled carbons to either the [1,4- ^{13}C]-succinic acid or the [2,3- ^{13}C]-succinic acid compound.

2.4 Conclusions

The 2D PDS spectra provide pronounced evidence for the preparation of biosynthetic site-specific pattern labeled LH2-samples. A relatively homogeneous label incorporation from the [1,4- ^{13}C] and the [2,3- ^{13}C] succinic acid precursor into the protein and BChl *a* cofactor of the LH2 complexes is demonstrated. Carbonyl

carbons from the membrane protein are mainly introduced from [1,4-¹³C]-succinic acid and the C α and C β are mainly labeled from [2,3-¹³C]-succinic acid with only a small fraction of label scrambling. In both the 1,4-LH2 and the 2,3-LH2 complexes the J-couplings are reduced, because in general the labeled C α carbons in the 2,3-LH2 are adjacent to an unlabeled CO carbon and the labeled CO carbons in the 1,4-LH2 are adjacent to an unlabeled C α carbon. The line width in the 2D spectra is about 1 ppm, which is larger than the effect of the J-coupling. The broadening can be due to a mild disorder in the LH2 sample. The labeled C β carbons only have an adjacent labeled C γ carbon in P, V, Q, E and K. Approximately each labeled C α has a labeled adjacent C β , which still form a pair of carbons with strongly coupled nuclear spins. The reduced labeling of the LH2 membrane protein conveniently reduces the spectral crowding and increases the resolution. This facilitates the identification of individual residues in comparison to uniformly labeled proteins. The specifically labeled 2,3-LH2 and 1,4-LH2 complexes provide a strong basis for a sequence specific assignment of the LH2 complex, which is described in chapter 3.

References

- (1) Egorova-Zachernyuk, T. A.; Hollander, J.; Fraser, N.; Gast, P.; Hoff, A. J.; Cogdell, R.; de Groot, H. J. M.; Baldus, M.; *J. Biomol. NMR* 2001, 19, 243-253.
- (2) McDermott, G.; Prince, S. M.; Freer, A. A.; Isaacs, N. W.; Papiz, M. Z.; Hawthornthwaite-Lawless, A. M.; Cogdell, R. J.; *Protein Eng.* 1995, 8, 43-43.
- (3) Prince, S. M.; Papiz, M. Z.; Freer, A. A.; McDermott, G.; Hawthornthwaite-Lawless, A. M.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 1997, 268, 412-423.
- (4) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 2003, 326, 1523-1538.
- (5) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H.; *Nature* 2002, 420, 98-102.
- (6) LeMaster, D. M.; Kushlan, D. M.; *J. Am. Chem. Soc.* 1996, 118, 9255-9264.

- (7) Heinen, W. In: *Grafting of polyolefins and miscibility in copolymer mixtures*; Leiden University: Leiden, 1996.
- (8) Hawthornthwaite-Lawless, A. M.; Cogdell, R. J. In: *The Chlorophylls*; Scheer, H., Ed.; CRC: Boca Raton, 1991; pp 494-528.
- (9) Metz, G.; Wu, X. L.; Smith, S. O.; *J. Magn. Reson. Ser. A* 1994, 110, 219-227.
- (10) Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W. G.; Lansbury, P. T.; Griffin, R. G.; *J. Chem. Phys.* 1998, 108, 9463-9479.
- (11) Castellani, F.; van Rossum, B. J.; Diehl, A.; Rehbein, K.; Oschkinat, H.; *Biochemistry* 2003, 42, 11476-11483.
- (12) Schulten, E. A. M.; Matysik, J.; Alia; Kiihne, S.; Raap, J.; Lugtenburg, J.; Gast, P.; Hoff, A. J.; de Groot, H. J. M.; *Biochemistry* 2002, 41, 8708-8717.

Chapter 3

Residual backbone and side-chain ^{13}C and ^{15}N resonance assignments of the LH2 complex*

3.1 Introduction

Insoluble and non-crystalline proteins, like most transmembrane proteins and protein aggregates, play a central role in intercellular communication and mediation of biological processes. This important class of proteins is involved in many biologically important functions, which can be understood thoroughly only when their detailed structures and interactions with the microscopic environment are known. Currently, only a few examples of transmembrane proteins with known crystal structure exist, which are mainly bacterial and plant photoreceptors.¹⁻³ Another example is rhodopsin, the G-protein coupled visual membrane photoreceptor, present in vertebrates, including the human eye.⁴ Generally, transmembrane proteins are not easily prepared in a crystalline or water-soluble form, because the lipids associated with these proteins impede crystallization

* The contents of this chapter have been accepted for publication in the *Journal of Biomolecular NMR*, 2005.

needed for X-ray crystallography and quench the rapid reorientation in solution which is a prerequisite for solution NMR spectroscopy. In contrast, for solid-state NMR spectroscopy local order and a homogeneous environment in the protein sample are sufficient. Hence, solid-state NMR has the potential of becoming a leading technology for structural investigations of transmembrane proteins.

In the last decade, Magic Angle Spinning (MAS) NMR has developed rapidly to resolve structures of microcrystalline peptide and protein samples at atomic resolution.⁵⁻¹⁰ These protein samples have β -sheet motifs containing polar residues. This leads to a favorable chemical shift dispersion and good spectral resolution, facilitating the chemical shift assignment of these preparations. In contrast, transmembrane proteins generally consist of α -helical segments containing many residues with aliphatic chain-chains, which are constrained all in virtually the same secondary structure, yielding a very narrow chemical shift dispersion. This complicates the sequence specific assignment of the many narrowly distributed and overlapping correlations.

On the basis of data analysis in many genome projects, it has been suggested that the relatively small membrane spanning proteins, which span the membrane 1-2 times, are most abundant of all transmembrane proteins in organisms.¹¹⁻¹³ These relative small transmembrane proteins contain about 30-70 residues. In this contribution, the 94 residues containing monomeric unit of the light-harvesting 2 (LH2) transmembrane protein complex from the anaerobic *Rhodospseudomonas (R.) acidophila* strain 10050 purple bacterium has been chosen as a model for the study of transmembrane proteins by MAS NMR. The monomer is a complex of two α -helical transmembrane segments, which are the α -subunit and β -subunit and the primary sequence is shown in figure 1.1. The LH2 complex is a good model for solid-state NMR technology development, because of the high abundance of this protein in the photosynthetic bacterium, which makes it convenient to introduce ¹³C and ¹⁵N isotope labels.^{3,14} In addition, the X-ray structure is known, which is helpful for this study. The LH2 system is a protein complex that contains cofactors, which can be also studied in the solid-state. Nine of these monomeric complexes form a concentric ring with nine-fold symmetry with a molecular weight of about 130 kDa. The complex is a part of a macroscopic assembly of many complexes, through which the light-harvesting is realized. Hence, primary, secondary, tertiary

and quaternary structures can be probed, as well as protein-cofactor arrangement and interactions.

At an early stage, the LH2 transmembrane protein complex has been used to explore the range and resolution of the MAS NMR in the study of transmembrane proteins.¹⁵ In the present study, biosynthetic ^{13}C isotope pattern labeling methods are used to overcome the spectral crowding, leading to the first sequence specific assignment of a real transmembrane protein by solid state NMR. A set of LH2 protein samples have been prepared, consisting of selectively labeled LH2 complexes obtained by a controlled growth of the bacteria in the presence of [1,2,3,4- ^{13}C]-succinic acid, [1,4- ^{13}C]-succinic acid, [2,3- ^{13}C]-succinic acid or a mixture of uniformly labeled amino acids.¹⁶ Referring to the isotopically labeled nutrient source in the expression medium these preparations are denoted as the U-LH2, 1,2,3,4-LH2, 1,4-LH2, 2,3-LH2 and AA-LH2 samples, respectively. The labeling patterns of all amino acid residues and for the Bacteriochlorophyll *a* (BChl *a*) cofactors in 1,2,3,4-LH2, 1,4-LH2, 2,3-LH2 are shown in figure 2.4.¹⁶ As shown in chapter 2, the biological growth conditions to prepare the AA-LH2 sample are different from the growth conditions for samples prepared from succinic acid. For succinic acid labeled samples, an excess of succinic acid was used to enhance the uptake of this nutrient source, while for the AA-LH2 sample preparation an excess of labeled amino acid mixture was used to enhance the uptake from the amino acid nutrient source. The scrambling pattern in the succinic acid labeled LH2 samples is different from the scrambling in the AA-LH2 sample. For the AA-LH2 sample, I and L are uniformly labeled, while also minor fractions of the A, G, V and P residues are labeled. By using both the samples labeled by succinic acid and the AA-LH2 sample, it is possible to arrive at a sequence specific assignment for 76 residues of the 94 in the LH2 complex, which is the largest membrane protein complex to date used for assignment studies with MAS NMR.

3.2 Materials and methods

For the NMR experiments, about 10 mg of U-LH2, 1,2,3,4-LH2, 1,4-LH2, 2,3-LH2 or AA-LH2 protein sample was transferred into a 4.0-mm CRAMPS rotor. On these

samples, 2D homonuclear ^{13}C - ^{13}C correlation spectra of labeled LH2 samples were recorded using PDS MAS NMR spectroscopy on a Bruker AV-750 spectrometer equipped with a double channel CP-MAS probe head and with a ^{13}C radio frequency of 188 MHz. The proton $\pi/2$ pulse was set to 3.1 μs , corresponding with a nutation frequency of 80.6 kHz. ^{13}C B_1 field strengths of 50 kHz corresponding with a cross polarization time of 2.0 ms were used during a 100-50% ramped CP sequence.¹⁷ In the PDS experiment, two-pulse phase modulation (TPPM) decoupling was applied during the t_1 and t_2 periods.¹⁸ A mixing time of 50 ms (PDS₅₀) was used to transfer the magnetization into the chain-chains, providing intra-residue correlations. Proton driven spin diffusion experiments with a long mixing time of 500 ms (PDS₅₀₀) have been applied to collect long-distance ^{13}C - ^{13}C inter-residue correlations.¹⁹

2D heteronuclear ^{13}C - ^{15}N correlation spectra were recorded with the same spectrometer using a triple resonance CP-MAS probe head and band selective NC magnetization transfer. ^{15}N polarization was created with a 80-100% ramped amplitude CP matching and a contact time of 2.0 ms. During ^{15}N evolution TPPM decoupling with a r.f. field strength of 81 kHz was used.^{17,18} To create a band-selective SPECIFIC-CP transfer from the nitrogen nuclei to either the CO or $\text{C}\alpha$, the carrier frequency was placed at 175 or 50 ppm, respectively.²⁰ During the dipolar contact time of 3.0 ms for the NCO transfer and 3.5 ms for the NCA transfer, a weak r.f. field of 22.5 kHz for ^{15}N was employed. The r.f. field strengths for ^{13}C were 14 kHz and 31 kHz for the NCA and NCO band selective CP, respectively. Band selectivity was achieved using adiabatic amplitude modulations on the ^{15}N channel.^{21,22} During the NC transfer off-resonance continuous wave decoupling was applied at the ^1H frequency.

For experiments involving homonuclear spin diffusion transfer, to obtain multiple correlations of multiple carbons with one nitrogen in a macromolecular network, a spin diffusion transfer period was included prior to the acquisition in t_2 . The $\pi/2$ pulses before and after the spin diffusion period were applied with a r.f. field of 45 kHz on the ^{13}C -channel. The spin diffusion period was 20 ms for the NCA(CO)CX and 30 ms for the NCOCACX transfer, where CX stands for any carbon atom. In both experiments, TPPM decoupling was applied during the t_1 and t_2 .¹⁸ All samples were cooled to 253 K, and the MAS spinning frequency $\omega_{\text{R}}/2\pi$ was 8.5 kHz (\pm 2Hz).

The ^{13}CO resonance of U- $^{13}\text{C},^{15}\text{N}$ -Tyrosine-HCl at 172.1 ppm was used as an external reference for the determination of the isotropic ^{13}C chemical shifts. The MAS NMR data were processed with XWINNMR software (Bruker) and subsequently analyzed using the program SPARKY (T. D. Goddard and D.G. Kneller, University of San Francisco)

3.3 Results and discussion

By applying PDSO and band selective SPECIFIC CP NMR methods on the 2,3-LH2, 1,2,3,4-LH2 and AA-LH2 preparations, an unambiguous sequence specific chemical shift assignment for backbone carbons and backbone nitrogens for the majority of the residues can be obtained. While in the past inter-residue correlations have been used to resolve 3-D structure, here they are used predominantly for a sequence specific assignment of the LH2 transmembrane protein complex.^{8,19}

The chemical shift assignments for residues from the α - and β -subunit are summarized in table 3.1a and table 3.1b at the end of this chapter, respectively. As a first step of the assignment procedure, the PDSO₅₀ spectrum of the U-LH2 in figure 3.1 was analyzed to identify the characteristic ^{13}C - ^{13}C correlation pattern of chain-chains of individual residues. For a major fraction of the residues characteristic chemical shift patterns could be observed. For 4 P, 8 T and 4 W residues, which occur both inside and outside the α -helix part, the number of spin systems in the spectra corresponds with the number of residues in the protein sequence. For the 13 A, 12 L and 9 V residues, abundantly present in the aliphatic α -helical segments of the transmembrane LH2 protein, the ^{13}C - ^{13}C correlations strongly overlap. For these residues it is difficult to resolve the spin systems corresponding with each residue. For instance, in the PDSO₅₀ spectrum three H spin systems and one K spin system have been observed, while H and K residues occur five and four times in the protein sequence, respectively. Hence two H spin systems and three K are either not detected or not resolved. Two of the five H residues and three of the four K residues are located in the flexible loops and termini of the protein outside the rigid α -helical part (figure 1.1). Since βK13 is

embedded in the rigid part of the α -helical segment of the β -subunit, the spin system of the K residue has been assigned to β K13. α H37 and β H41 are in the flexible part of the protein and most likely their correlation signals are quenched by dynamics. The three observable H spin systems are attributed to α H31, β H12 and β H30. In two of these H spin systems the $C\alpha$ and $C\beta$ responses are shifted upfield. They are assigned to α H31 and β H30, since these residues coordinate to the central Mg ion in the conjugated macrocycles of the BChl cofactors, which induce an upfield ring current shift.^{23,24} The third set has been assigned to β H12, which is in the rigid α -helical segment of the β -subunit.

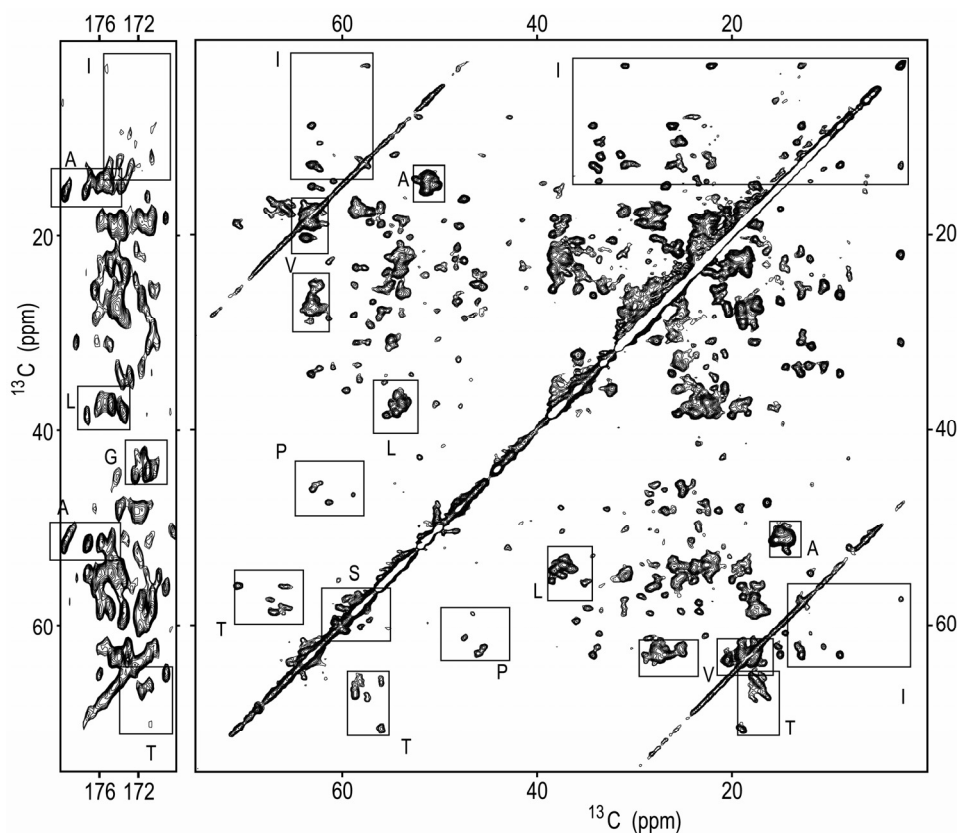


Figure 3.1 The aliphatic part of the PDSD₅₀ ^{13}C - ^{13}C correlation spectrum of the U-LH2. Most type of amino acid residues can be identified. Correlations of I, L, T, P, S, A and V residues are indicated with rectangles.

By using the 1,2,3,4-LH2, 2,3-LH2 and AA-LH2 samples, the overlap in the PDS₅₀ data of the U-LH2 is reduced. According to figure 2.4, the majority of the C α and C β carbons of the residues synthesized from succinic acid are ^{13}C labeled by using [2,3- ^{13}C]-succinic acid. The aliphatic carbons of the H residues are not labeled in the 2,3-LH2 sample and the C α -C β signals of this residue are assigned from the PDS₅₀ dataset collected from 1,2,3,4-LH2. For the other residues, labeled by succinic acid, the PDS₅₀ spectrum of 2,3-LH2 has been used to assign C α -C β correlations. For detection of the I and L residues, the AA-LH2 was used. Due to the uptake of labels from the amino acid mixture, a minor fraction of the A, G, V and P residues is also labeled in this sample, and the AA-sample has been used to improve the assignments for these residues.

In the upper right panels in figures 3.2 (p. 42) and 3.3 (p. 43), the PDS₅₀ spectrum of 2,3-LH2 and the PDS₅₀ spectrum of AA-LH2 are shown. Compared to the PDS₅₀ spectrum of the U-LH2 in figure 3.1, the number of correlations is reduced, while the resolution in the spectra is improved. For backbone carbons, there is a significant reduction of J-couplings, which contributes to the resolution improvement. For most side-chain carbons the reduction of the scalar coupling is only small, which provides only a marginal improvement of the resolution. The data from the pattern labeled samples shows also a reduced number of correlations in the aliphatic region between 0 and 75 ppm, since the I and L residues, which represent about 20% of the total protein, are not enriched when growing from the labeled succinic acid.

By aligning the PDS₅₀ spectrum of 2,3-LH2 with the ^{15}N - ^{13}C NCACX and NCACOCX correlation spectra of various pattern labeled samples, all C α -C β correlations in the spectrum can be assigned to specific residues. The eight T and four S residues are rapidly identified due to the relative downfield shifts of both T/S(C α) and T/S(C β) chemical shifts. To illustrate the assignment procedure, correlation responses of the $\alpha\text{T}38$ spin system are indicated in figure 3.2. The C α -C β correlations of T and S residues are well resolved between 56 and 71 ppm, close to the diagonal. The T(C α -C γ) and T(C β -C γ) correlations, observed in the spectrum of the U-LH2 are quenched in the spectrum of 2,3-LH2, which shows that T(C γ) is not labeled. The PDS₅₀ spectra also show unique P(C α -C δ) correlations, which is important for the assignment of the P residues, shown in figure 3.2.

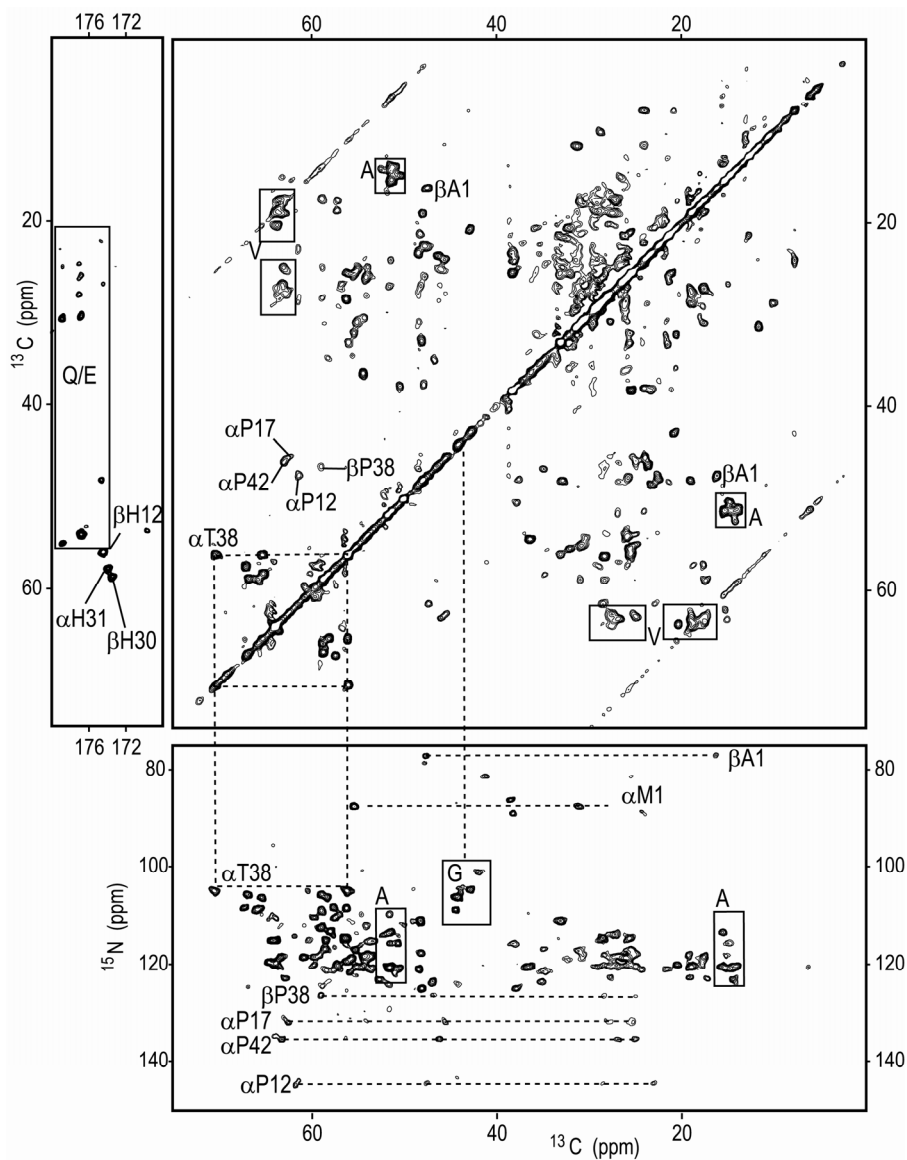


Figure 3.2 In the upper part of the figure the ^{13}C - ^{13}C PDS₅₀ correlation spectrum collected from 2,3-LH2 is shown. The region shown in the upper left panel contains responses involving the aliphatic carbons and carbonyl carbons, while the right panel shows correlations between aliphatic carbons. In the left part, a few responses are observed for 2,3-LH2, belonging to H, Q and E residues. The correlations involving the αT38 and four P residues for the 2,3-LH2 are indicated. Correlations for A, G, M and V residues are also indicated. The correlations can be correlated with the aliphatic region of the NCACX spectrum shown at the bottom.

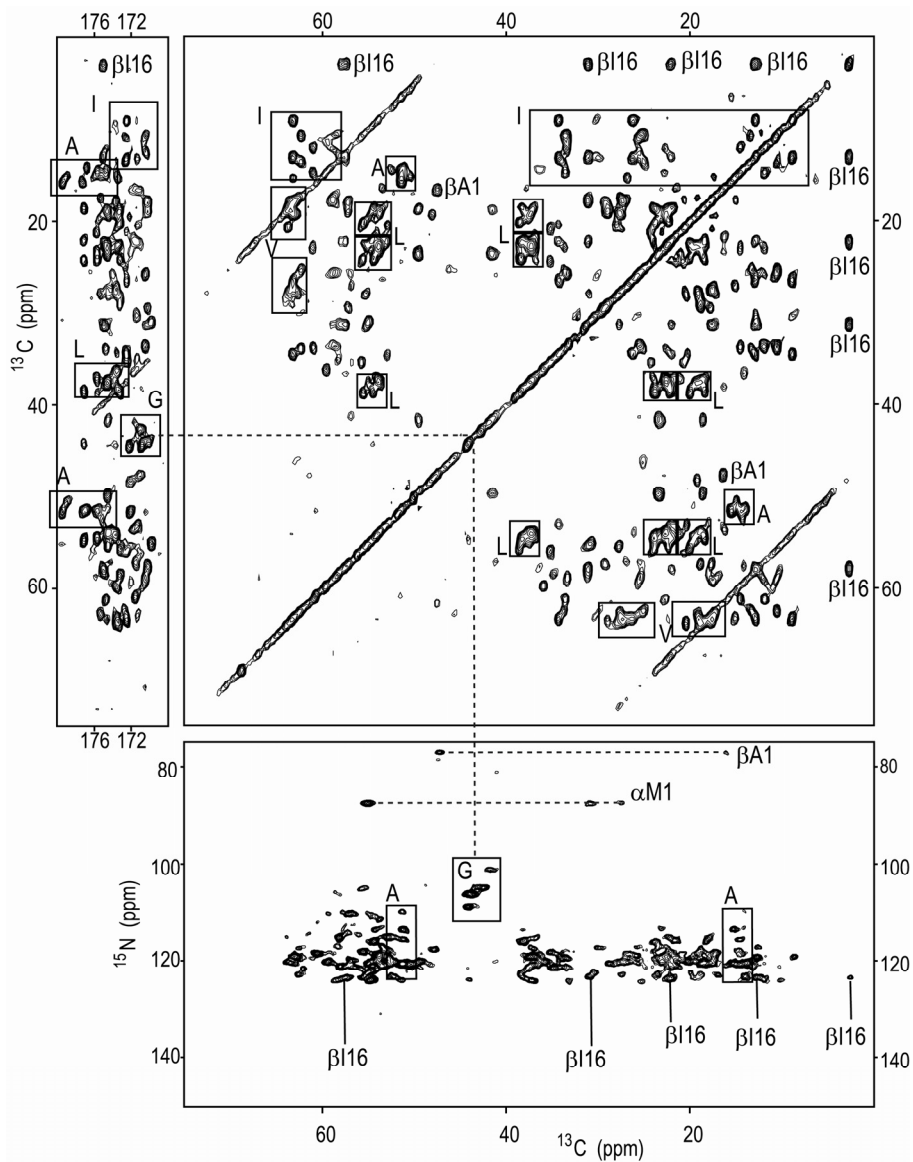


Figure 3.3 In the upper part of the figure the ^{13}C - ^{13}C PDSD₅₀ correlation spectrum obtained on AA-LH2 is shown. The region shown in the upper left panel contains cross peaks involving the aliphatic carbons and carbonyl carbons, while the right panel shows correlations between aliphatic carbons. In this spectrum the I, L, V, A and G residues are clearly observed. In the panel at the bottom the correlations from the PDSD₅₀ spectrum can be correlated with the aliphatic region of the NCACX spectrum. βI16 is indicated as a clear example of a chemical shift data set.

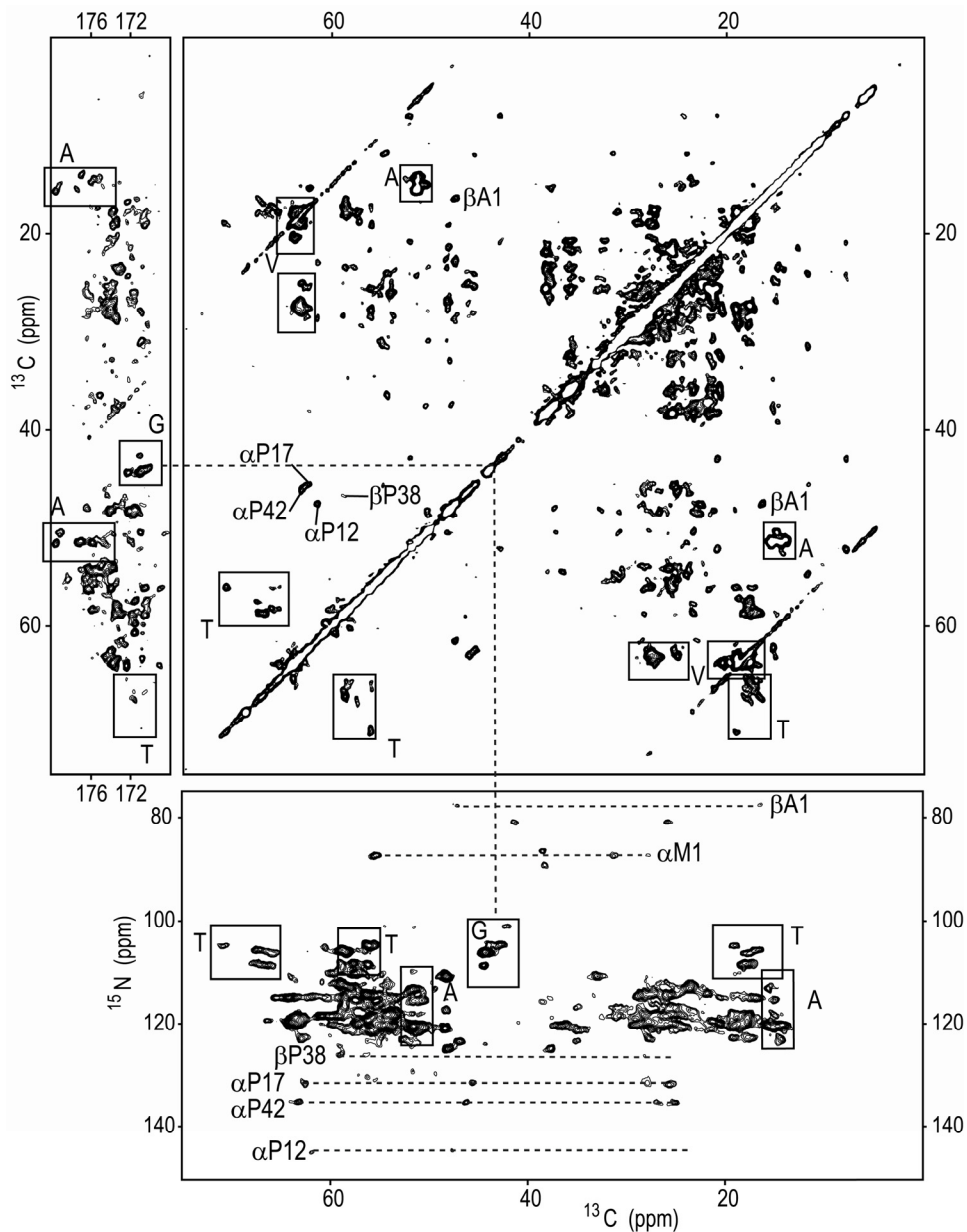


Figure 3.4 The ^{13}C - ^{13}C PDSD₅₀ correlation spectrum and NCACX spectrum (bottom) obtained on 1,2,3,4-LH2. Correlations for most of the residues can be observed. Because the I and L residues are quenched, the figure clearly shows that these residues are not synthesized from succinic acid.

As an example of the responses of the I residues from the AA-LH2 sample, the signals from βI16 , located at the upfield limit of the aliphatic spectrum are indicated in figure 3.3. In this way the characteristic multiple-bond correlation signal pattern for this type of residue is illustrated. Corresponding carbon correlation networks are observed for αI6 , αI14 , αI16 , αI26 , αI28 and αI34 . In contrast, the correlations of the L residues strongly overlap and are difficult to resolve. The A and V residues are observed in both the 2,3-LH2 and AA-LH2 dataset (figures 3.2 and 3.3, respectively). In the U-LH2 dataset V($\text{C}\alpha\text{-C}\beta$) correlations overlap with I ($\text{C}\alpha\text{-C}\gamma$) and can be resolved by using the spectra from the pattern labeled samples.

The CO- $\text{C}\alpha\text{-C}\beta$ correlation area for the 2,3-LH2 shows correlation signals for E, Q and H residues, which are the only residues that have both a ^{13}CO and a $^{13}\text{C}\alpha$ backbone carbon. To observe the CO- $\text{C}\alpha\text{-C}\beta$ correlations for other residues synthesized from succinic acid, the spectrum of the 1,2,3,4-LH2 in the upper left panel of figure 3.4 (p. 44) can be used, since in the 1,2,3,4-LH2 both the CO and the $\text{C}\alpha$ are labeled. For the assignment of the chemical shifts for the I, L and V residues in this spectrum, in particular for the crowded CO- $\text{C}\alpha\text{-C}\beta$ correlation area, the simplification due to the pattern labeling is essential. Assignments for G residues, which have no aliphatic chain-chains, are obtained starting from the G($\text{C}\alpha$) response between 40-43 ppm in the CO- $\text{C}\alpha$ correlation region, where 5 out of 7 G residues are identified. Two residues, αG47 and αG48 , are in the mobile part near the C-terminus of the α -subunit and are broadened beyond the limit of detection.

The ^{13}C chemical shift correlation patterns in the PDSD₅₀ spectra can be aligned with the $^{15}\text{N}\text{-}^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta\text{-}^{13}\text{C}\gamma\text{-}(^{13}\text{C}\delta)$ (NCACX) correlation spectra to resolve the correlations with the ^{15}N backbone signals. Previously, it has been shown that the $^{15}\text{N}\text{-}^{13}\text{C}$ correlation spectra of U-LH2 are too crowded to extract the individual chemical shifts.¹⁵ Figures 3.2, 3.3 and 3.4 also present the NCACX spectra of 2,3-LH2, AA-LH2 and 1,2,3,4-LH2, respectively. It is clear that the use of the various pattern-labeled samples separates many responses that overlap in the data collected from the U-LH2 (figure 3.1). Starting with the NCACX spectrum of 2,3-LH2 in figure 3.2, intra-residual $^{15}\text{N}\text{-}^{13}\text{C}\alpha\text{-}^{13}\text{C}\beta\text{-}(\text{C}\gamma)$ correlation networks are resolved and can be assigned by alignment with the correlation sets collected from the PDSD spectrum of this sample. For instance, $\alpha\text{T38}(\text{C}\gamma)$ clearly shows up in the PDSD₅₀

spectrum of 1,2,3,4-LH2, but not in the PDSD₅₀ spectrum of 2,3-LH2, because the C γ is unlabeled (figure 2.4). The ¹⁵N chemical shifts of the α P12, α P17, α P42 and β P38, i.e. all 4 P residues of the LH2 complex, resonate downfield with respect to the ¹⁵N response of other residues. ¹⁵N responses for G residues are clearly resolved in the NCACX data from their N-C α correlations that are relatively upfield on both the ¹³C and ¹⁵N chemical shift scales, and align well with the G(CO-C α) responses in the PDSD₅₀ spectrum.

The NCACX spectrum of AA-LH2 in the bottom panel of figure 3.3 shows ¹⁵N-¹³C correlations from labeled I and L residues. In addition, V, A and G residues are observed due to the partial incorporation of labels from the amino acid mixture nutrient source. The correlations involving these aliphatic residues in the α -helical part predominantly show up with ¹⁵N chemical shifts of ~120 ppm. The ¹⁵N shifts of the I residues have been assigned on the basis of the ¹⁵N-C δ correlations between 2 and 18 ppm ¹³C shift. For example, the β I16 response in the NCACX spectrum is well in line with the corresponding correlation pattern in the PDSD₅₀ spectrum shown in the upper right panel of figure 3.3. The upfield shifted C δ response of β I16 at 2.5 ppm is due to the short distance of 3.7 Å between this nucleus and the aromatic macrocycle of the B800 BChl cofactor, which produces a significant ring current shift.²⁵ For α M1 at the N-terminus of the α -subunit, a correlation set is found at 87.6 ppm ¹⁵N shift. The carboxyl- α M1-N-terminus is coordinated to the central Mg-ion of the B800 cofactor, which contributes to the rigidity of the N-terminal loop of the α -subunit. This accounts for the observation of many residues outside the α -helical segments in our data.

The majority of the ¹⁵N-¹³C correlations of the V residues coincide with ¹⁵N-¹³C signals from I and L residues. In the NCACX data of the 2,3-LH2 sample the signals from the V residues are clearly resolved since the responses of the I and L residues are not present. A residues can be identified in the PDSD₅₀ spectra of both the AA-LH2 and the 2,3-LH2. The C α -C β correlation signals of the A residues strongly overlap in the PDSD₅₀ spectra. A(C α) shifts are observed between 47 and 53 ppm, while A(C β) signals are observed between 14 and 16 ppm. To distinguish between the various A residues the CO-C α -C β correlation area is used, since the dispersion of the CO responses is largest. In the NCACX spectra, A residues can be

distinguished conveniently because the A(C α) signals are shifted upfield, while the ^{15}N backbone signals are between 117 and 124 ppm. One exception is the ^{15}N response at 77.2 ppm, which correlates with C α and C β responses at 47.7 and 16.6 ppm, respectively. These correlations have been assigned tentatively to the $\beta\text{A}1$, which is at the N-terminus of the β -subunit.

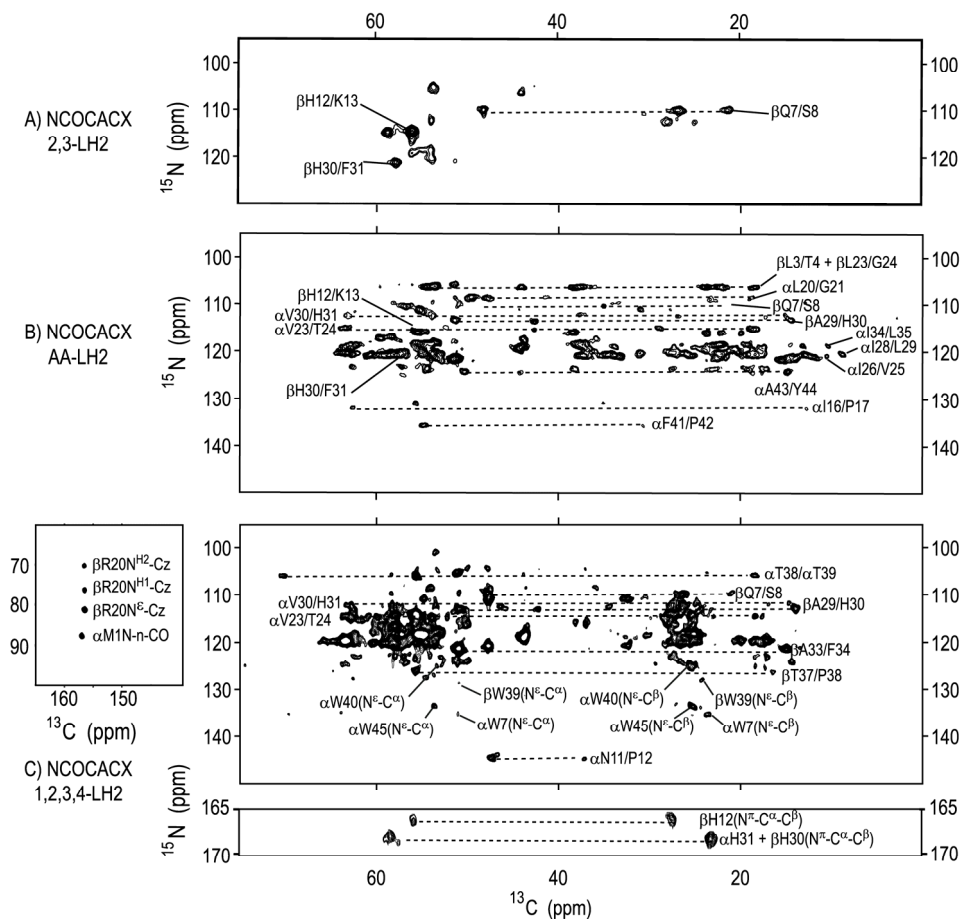


Figure 3.5 The aliphatic region of the NCOACAX spectra of 2,3-LH2 (A), AA-LH2 (B) and 1,2,3,4-LH2 (C). Several sequence specific assignments are indicated in the figure.

For a sequence specific assignment of the spin systems inter-residual magnetization transfer from the backbone N_i to the $C_{i-1}O$ carbon was produced using the band selective SPECIFIC CP method.²⁰ The N_i - CO_{i-1} cross polarization, where the magnetization is transferred into the chain-chain of a previous residue, yields $N_iCO_{i-1}CA_{i-1}CX_{i-1}$ (NCOCACX) correlation spectra that can be aligned to both the PDS_{D50} and the NCACX spectra to identify the correlating residues.

Figure 3.5A, 3.5B and 3.5C shows the NCOCACX correlation spectra of 2,3-LH2, AA-LH2 and 1,2,3,4-LH2, respectively. For observation of the residues in the aliphatic part of the NCOCACX spectra, both the CO_i and the Ca_i positions should be ^{13}C labeled and covalently connected with $^{15}N_{i+1}$. Such labeling patterns occur for H, Q and E residues in the 2,3-LH2. The correlations of the 2,3-LH2 in the NCOCACX spectrum can be aligned with the correlations in the CO region of the PDS_D spectrum in the upper left panel of figure 3.2. In the NCOCACX spectrum of 2,3-LH2 in figure 3.5A, 2 out of 5 H residues are observed, while in the PDS_D spectrum 3 out of 5 H-residues are detected. The protein sequence shows that $\alpha H31$ is followed by $\alpha L32$, which is unlabeled in this sample. The 2 inter-residual $H(C\alpha_i-N_{i+1})$ correlations in the 2,3-LH2 NCOCACX spectrum can now be assigned unambiguously to $\beta H30/\beta F31$ and $\beta H12/\beta K13$, while the third H correlation set, which is only present in the PDS_{D50} spectrum, can be assigned to $\alpha H30$. This assignment is well in line with data collected from LH2 samples where only H residues are labeled.²⁶ Two spin systems for the Q and E responses in this spectrum are assigned to $\beta Q7$ and $\beta E10$. $\beta Q7$ can be assigned due to the correlation with the nitrogen of $\beta S8$. Since $\beta E10$ is followed by the unlabeled $\beta L11$ in the 2,3-LH2, a correlation with the $\beta E10$ is not detected in the NCOCACX spectrum of this sample. This leads to the specific assignment of the $\beta E10$ responses.

The NCOCACX spectrum of the AA-LH2 in figure 3.5B shows strong I_i-I_{i+1} , L_i-L_{i+1} , A_i-A_{i+1} , V_i-V_{i+1} and G_i-G_{i+1} correlations. Like in the NCACX spectrum, most ^{15}N - ^{13}C correlations in the NCOCACX spectrum of AA-LH2 have ^{15}N chemical shifts close to 120 ppm that are difficult to resolve. Some of the correlation sets that are not resolved in the NCACX spectrum can be assigned using the upfield or downfield shifted ^{15}N response of the next residue. Examples are $\alpha L20/G21$, $\alpha V23/T24$, $\alpha V30/H31$, $\alpha I16/P17$, $\alpha F41/P42$, $\beta A29/H30$, $\beta L3/T4$ and $\beta L23/G24$ (figure 3.5B). Inter-residue $^{15}N_{i+1}$ - $^{13}C\delta_i$ correlations are observed between 2 and 18 ppm ^{13}C shift

for $\alpha\text{I24/L35}$, $\alpha\text{I28/L29}$ and $\alpha\text{I26/LV25}$. They can be assigned by aligning the spectrum to the NCACX spectrum of the AA-LH2 in figure 3.3.

To observe inter-residual correlations of other residues synthesized by the succinic acid nutrient source, the NCOCACX spectrum of 1,2,3,4-LH2 is used. Residues synthesized from [1,2,3,4- ^{13}C]-succinic acid are labeled on both the $\text{C}\alpha$ and the CO position. The NCOCACX spectrum of 1,2,3,4-LH2 in figure 3.5C shows a strong reduction of correlation signals with respect to the NCACX spectrum in figure 3.4. Since the I and L residues are unlabeled in the 1,2,3,4-LH2, the $\text{I}_i\text{-I}_{i+1}$ and $\text{L}_i\text{-L}_{i+1}$ responses are not present in the NCOCACX spectrum. This reduces the spectral crowding by about 40% with respect to the NCOCACX spectrum of U-LH2. Hence ^{13}C correlations sets that are observed in the NCACX spectrum of the 1,2,3,4-LH2 and are not observed in the NCOCACX spectrum, can be assigned to residues that are followed by an I or L residue in the sequence. Labeled residues that are not followed by a L or I in the protein sequence give rise to correlations in the NCOCACX spectra of 1,2,3,4-LH2 and AA-LH2. Clear examples are the $\alpha\text{V30/H31}$ and $\beta\text{A29/H30}$ correlations. These two examples also demonstrate how two nearby correlations of the $\alpha\text{H31}(\text{C}\alpha/\text{C}\beta)$ and $\beta\text{H30}(\text{C}\alpha/\text{C}\beta)$ in the PDSD₅₀ spectra can be distinguished with help of the correlation to the adjacent residue, leading to an unambiguous assignment. The inter-residue correlations of $\alpha\text{T38/T39}$, $\beta\text{T37/P38}$ and $\alpha\text{N11/P12}$ are only present in the NCOCACX data collected from the 1,2,3,4-LH2, since both the $\text{C}\alpha$ and the CO carbons of these residues are labeled in this sample.

The NCOCACX of the 1,2,3,4-LH2 sample also reveal intra-residue correlations between chain-chain nitrogens and carbons of the W, R and H residues. The ^{15}N chemical shifts from the chain-chain nitrogens are significantly different from the backbone ^{15}N shifts and can be conveniently used for the assignment of residues, facilitating also the sequence specific assignment. For instance in figure 3.5C, the $\text{N}\epsilon\text{-C}\alpha\text{-C}\beta$ correlations for αW7 , αW40 , αW45 and βW39 are indicated. In addition, a set of three intense correlations are observed for the $\text{C}\gamma$ -carbon of βR20 at 156.2 ppm with $\text{N}\epsilon$, NH1 and NH2 at 81.2, 76.2 and 70.1 ppm. Also three sets of $\text{N}\pi\text{-C}\gamma\text{-C}\delta\text{-C}\beta\text{-C}\alpha$ correlations at $^{15}\text{N} = 168.1, 166.4$ and 169.0 ppm are observed for αH31 , βH12 and βH30 , respectively.

To collect more inter-residue correlations, PDS_{D500} data were collected from the 2,3-LH2 and AA-LH2 samples. The reduction of ¹³C labeling in these samples strongly suppresses relayed spin diffusion. This enables the magnetization transfer over longer distances than for uniformly labeled samples and provides the possibility to detect correlations between remote carbons from different residues. In contrast to the band selective magnetization transfer in the NCACOCX method, which yields selectively *i*/*i*-1 correlations, the magnetization in the PDS_{D500} method is transferred to any nearby ¹³C nuclei. The PDS_{D500} spectra obtained from the 2,3-LH2 and AA-LH2 are shown in figure 3.6 and 3.7 on p. 51 and p. 53, respectively. Using the intra-residue assignments of the PDS_{D50} spectrum as a starting point, the long-range correlations in the PDS_{D500} spectrum can be assigned to combinations of residues that are unique in the protein sequence. In this way, the long-range correlations contribute to a sequence specific assignment of the protein NMR response.

In the PDS_{D500} spectrum of the 2,3-LH2 in figure 3.6, a clear example of long-range inter-residue correlations is provided for α Thr38. The α Thr38 is in the middle of a short loop, which connects the two α -helical segments in the α -subunit. In addition to the α Thr38(C α -C β) correlation at 56.1-70.6 ppm, the C β signal of this residue at 70.6 ppm correlates with a unique set of signals, which can be assigned unambiguously to α P42(C δ), α A43(C α), α W40(C α), α W40(C β), α S36(C α) and α T39(C α). The cross-peaks of the α H37 ¹³C responses with α T38(C β) are attenuated in the PDS_{D500} spectrum. This is attributed to the mobility of α H37, and is in line with the absence of its signal in the PDS_{D50} spectra. A second example of long-range interactions is β F22. In the protein sequence, β F22 is located between labeled β V21 and unlabeled β L23. In line with the labeling pattern, only the β F22/V21 nearest neighbor signals are detected. In addition, β F22/ β G24 correlations are weakly observed.

Since [2,3-¹³C]-succinic acid is an important precursor for the biosynthesis of BChls in photosynthetic bacteria, the BChls in the 2,3-LH2 are also pattern labeled by following the biosynthetic [2,3-¹³C]-succinic acid or [1,4-¹³C]-succinic acid incorporation pathway.²⁷ For 2,3-LH2 most of the peripheral groups of the BChls that interact with the protein matrix are labeled and the PDS_{D500} spectrum of 2,3-LH2 in figure 3.6 clearly shows α C12¹/ β V28/ β A29/ β H30 and

The additional correlations in the PDS_{D500} spectrum of the AA-LH2 compared to the PDS_{D50} dataset collected from the same sample can be attributed to long-range correlations between remote I, L, G, A, and V residues. A clear example in the spectrum is provided for the β I16 in figure 3.7, where β I16(C δ) correlates with the ¹³C responses of the β V15, which is the only labeled residue in the vicinity of β I16. Also the α I26/A27, α I28/A27/L29 and α I34/A33/L35 give rise to long-range correlation signals. These long-range correlations can be related to unique combinations in the protein sequence and assigned unambiguously. For the α G15/I16 α G21/ α L19 and β G24/ β L23/ β L25 long-range correlations in the PDS_{D500} spectrum between I and L with G help to resolve signals from these aliphatic residues, which are badly resolved in the PDS_{D500} spectrum due to strong overlap. Finally, signals of the A residues that strongly overlap in the PDS_{D50} spectra of 2,3-LH2 and AA-LH2 are also conveniently resolved in the PDS_{D500} by remote correlations with the A(C α) and A(C β) carbons.

Most of the long-range correlations in both PDS_{D500} spectra in figure 3.6 and 3.7 are between residues that are within one α -helical turn. The strongest long-range correlations are observed for adjacent residues ($i, i\pm 1$) in the protein sequence, while the intensities of the correlation signals for $i, i\pm 2$ and $i, i\pm 3$ are much weaker. The inter-helix distance in the LH2 complex of about 9.0-10 Å is too large for the observation of inter-helical correlations between polypeptide backbone atoms. This contrasts with β -strand proteins. Here inter-strand correlations are rapidly observed since the distances between two adjacent β -strands are in the range from 4.6-5.4 Å.^{8,19} It is known that interhelical side-chain side-chain contacts can be as short as 4.5 Å. Carbons involved in these contacts are predominantly aliphatic and resonate between 10 and 20 ppm. Since there are also many inter-residue correlations, it is difficult to discriminate the long-range from the sequential correlations in this small and crowded part of the correlation spectrum with many correlations that are close to the diagonal.

increases the spectral resolution. In particular for backbone carbons the J-couplings in the samples are reduced significantly. It is demonstrated that the use of pattern labeled samples subdivides the many overlapping correlations of the uniformly labeled transmembrane protein samples into data sets with resolved correlations and sequence specific assignments are obtained. The dilute isotope enrichment also paves the way for long-distance transfer of the magnetization, enabling the detection of inter-residue correlations that are used for the sequence specific assignment and for obtaining information about the structure of the protein complex.

The solid-state assignment of the LH2 complex is nearly complete, albeit that the signals of the residues α 47-53 and β 40-41 at the two C-termini are not detected in the MAS NMR. This indicates that the C-termini are either flexible or disordered in the protein samples. In contrast, N-termini are clearly observed, showing that the N-termini are ordered uniformly. The order for the α N-terminus is possibly a result of the non-covalent bond between the N-carboxyl- α M1 and the cofactor. The β N-terminus, consisting of only the first four residues, is very short which may explain why it is not significantly disordered. The structure of the rigid part is unique, since no doubling of the resonances is observed. The chemical shift information obtained from the spectra can be used to predict the secondary structure of the LH2 protein in a future step.

In addition, unique correlations are observed between carbons from the BChl α cofactor and the LH2 protein, which provide structural information of the active site and useful structural information of the organization of the subunits, which interact with the cofactor, i.e. the quaternary structure of the protein. The sequence specific assignment of the LH2 complex in this study shows that it is possible to assign MAS NMR data from long α -helical segments in transmembrane proteins, and paves the way for structure determination of this important class of biological constructs.

Table 3.1a. Solid-state ^{15}N and ^{13}C chemical shift assignments for the α -subunit of the LH2 transmembrane protein complex. These chemical shifts have been deposited with the BioMagResBank (BMRB) under the accession number 6348.

Residue	^{15}N	^{13}CO	$^{13}\text{C}^\alpha$	$^{13}\text{C}^\beta$	$^{13}\text{C}^\gamma$	$^{13}\text{C}^\delta$	$^{13}\text{C}^\epsilon$	$^{15}\text{N}^\epsilon$	$^{15}\text{N}^\zeta$
αM1	87.6	169.8	55.3	31.2	27.8				
αN2	123.5	172.4	46.8	34.9					
$[\alpha\text{Q3}]$	112.8		53.7						
$[\alpha\text{G4}]$	104.7	170.9	42.6						
αK5	113.6								
$[\alpha\text{I6}]^a$		174.0	57.6	36.5	14.1/25.3	9.0			
αW7	[121.0]	172.0	51.4	24.1	109.1	125.5	137.7	135.6	
αT8	110.2	174.4	56.7	64.1	17.8				
αV9									
αV10	120.5	170.3	57.2	29.2	17.7/19.0				
αN11	125.0	173.6	48.0	37.6	176.3				
αP12	144.8	172.9	61.5	28.6	22.9	47.5			
αA13	115.3	177.1	51.5	14.7					
$[\alpha\text{I14}]^a$	119.7	174.6	61.1	33.6	14.5/22.6	12.0			
αG15	105.0	170.1	43.9						
αI16	117.3	171.4	62.9	30.2	24.0/12.7	8.8			
αP17	131.7	173.9	62.5	28.2	25.4	45.5			
$[\alpha\text{A18}]^b$			51.3	14.9					
$[\alpha\text{L19}]^b$				37.2	22.0	20.0/18.5			
αL20	120.2	174.5	49.6	41.5	23.6	23.6/18.6			
αG21	108.8	172.1	44.3						
αS22	118.8	171.5	59.8	60.7					
αV23	120.6	173.3	63.9	29.1	18.3/19.1				
αT24	115.0		64.1	65.2	16.7				
αV25	122.9	175.3	62.7		17.6/19.1				
αI26	124.0	170.4	58.4	33.4	25.2/12.3	10.45			
αA27	120.8	177.2	51.5	15.3					
αI28	119.3	172.4	63.3	34.3	26.3/13.0	9.0			
αL29	120.5		54.3	37.5	22.6	17.7/19.1			
αV30	118.3	173.5	63.2	26.9	15.1/17.4				
αH31	112.4	173.0	58.5	23.4	125.1	118.2	132.9		168.1
$[\alpha\text{L32}]^b$		173.7	53.8	38.0	23.2	16.5/21.6			
$[\alpha\text{A33}]^b$	120.4	176.4	51.6	13.2					
αI34	121.4	172.9	62.4	33.7	25.3/13.3	10.6			
αL35	119.0		54.4	37.9	22.7	17.8/20.6			
αS36	114.9	173.1	58.3	59.8					
αH37	not observed								
αT38	105.0	170.6	56.1	70.6	19.0				
αT39	107.2	169.0	56.0	64.1	17.1				
αW40	[118.6]	173.7	53.5	25.9	108.9	119.7	134.9	125.2	
αF41	122.4	169.9	54.9	30.8	134.7				
αP42	135.3	174.5	63.0	28.0	25.0	46.1			
αA43	120.8	178.8	50.4	14.9					
αY44	124.1	177.2	51.6	44.2					
$[\alpha\text{W45}]^a$	118.6	173.6	54.0	25.8	105.3	126.1	136.3	133.7	
$\alpha\text{Q46-}\alpha\text{A53}$	not observed								

Tentative assignments are given in brackets.

a) $\alpha\text{W45}/\beta\text{W39}$ and $\alpha\text{I6}/\alpha\text{I14}$ data sets can be interchanged.

b) $\alpha\text{A18}/\alpha\text{L19}$, $\alpha\text{L32}/\alpha\text{A33}$ and $\beta\text{L25}/\beta\text{A26}/\beta\text{L2}$ spin systems are assigned from the PDS₅₀₀ of the AA-LH2 sample. Because of the frequent AL or LA combination in the sequence, a sequence specific assignment is difficult.

Table 3.1b. Solid-state ^{15}N and ^{13}C chemical shift assignments for the β -subunit of the LH2 transmembrane protein complex. These chemical shifts have been deposited with the BioMagResBank (BMRB) under the accession number 6348

Residue	^{15}N	^{13}CO	$^{13}\text{C}^{\alpha}$	$^{13}\text{C}^{\beta}$	$^{13}\text{C}^{\gamma}$	$^{13}\text{C}^{\delta}$	$^{13}\text{C}^{\epsilon}$	$^{15}\text{N}^{\epsilon}$	$^{15}\text{N}^{\zeta}$
[β A1]	77.1	171.2	47.5	16.3					
β T2	108.9	170.3	58.8	66.0	17.1				
β L3	115.9	177.1	54.4	38.3	24.1	18.4/22.1			
β T4	106.1	170.6	58.7	66.8	16.8				
β A5	118.0		51.5	14.3					
β E6	118.1	176.1	54.1	25.9	30.5				
β Q7	120.9	174.0	48.2	21.8	26.8				
β S8	110.2	173.6	57.3	59.6					
β E9									
β E10	117.2	178.3	55.1	25.2	30.8				
β L11									
β H12	119.4	173.8	56.0	27.6	132.9	113.5	133.7		166.4
[β K13]	114.5	170.7	50.4	31.0	23.4	27.8	48.4		
β Y14									
β V15	114.2		62.1	27.7	17.1/19.0				
β I16	123.6	174.8	57.6	31.1	22.1/12.9	2.7			
β D17	115.6	173.6	50.0	37.6					
β G18	105.5	170.8	44.2						
β T19	106.4	176.0	58.0	65.8	17.5				
β R20 ^c		171.4	56.9	26.8	25.5	41.1		81.2	
β V21									
β F22	120.6		54.4	36.2					
β L23		174.5	53.9	37.4	23.9	18.8/22.9			
β G24	106.3	171.9	44.0						
[β L25] ^b	119.2		54.0	37.4	23.6	19.5/22.1			
[β A26] ^b	120.8		50.9	13.9					
[β L27] ^b			54.1	37.5					
β V28	120.2	173.1	63.7	27.4	17.6/20.5				
β A29	120.3	176.1	51.4	14.4					
β H30	113.1	173.2	57.8	23.4	125.2	118.8	133.5		169.0
β F31	120.7	173.5	59.6	36.0	136.2				
[β L32]	117.4		53.7	38.3	22.6	19.1/20.4			
[β A33]	120.51		51.5	15.5					
β F34	121.5	175.8	58.3	34.9	134.7				
β S35	105.0		55.9	59.7					
β A36	120.5	173.8	50.5	16.8					
β T37	108.6	170.5	57.3	67.4	17.3				
β P38	126.5		58.9	28.4	25.2	46.6			
[β W39] ^a	120.53		55.0	25.0	108.5	123.0	135.4	128.2	
β L40- β H41	not observed								

Tentative assignments are given in brackets.

- α W45/ β W39 and α I6/ α I14 data sets can be interchanged.
- α A18/ α L19, α L32/ α A33 and β L25/ β A26/ β L2 spin systems are assigned from the PDS₅₀₀ of the AA-LH2 sample. Because of the frequent AL or LA combination in the sequence, a sequence specific assignment is difficult.
- The chemical shifts for C^{ζ} , NH^1 , NH^2 of β R20 are 156.2, 76.2 and 70.1 ppm, respectively.

References

- (1) Koepke, J.; Hu, X. C.; Muenke, C.; Schulten, K.; Michel, H.; *Structure* 1996, 4, 581-597.
- (2) Kuhlbrandt, W.; *Nature* 2001, 411, 896-899.
- (3) Prince, S. M.; Papiz, M. Z.; Freer, A. A.; McDermott, G.; HawthornthwaiteLawless, A. M.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 1997, 268, 412-423.
- (4) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M.; *Science* 2000, 289, 739-745.
- (5) Bockmann, A.; Lange, A.; Galinier, A.; Luca, S.; Giraud, N.; Juy, M.; Heise, H.; Montserret, R.; Penin, F.; Baldus, M.; *J. Biomol. NMR* 2003, 27, 323-339.
- (6) Hong, M.; *Biophys. J.* 1999, 76, A392-A392.
- (7) Detken, A.; Hardy, E. H.; Ernst, M.; Kainosho, M.; Kawakami, T.; Aimoto, S.; Meier, B. H.; *J. Biomol. NMR* 2001, 20, 203-221.
- (8) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H.; *Nature* 2002, 420, 98-102.
- (9) Pauli, J.; Baldus, M.; van Rossum, B.; de Groot, H.; Oschkinat, H.; *Chembiochem* 2001, 2, 272-281.
- (10) Straus, S. K.; Bremi, T.; Ernst, R. R.; *J. Biomol. NMR* 1998, 12, 39-50.
- (11) Wallin, E.; von Heijne, G.; *Protein Sci.* 1998, 7, 1029-1038.
- (12) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. L.; *J. Mol. Biol.* 2001, 305, 567-580.
- (13) Simon, I.; Fiser, A.; Tusnady, G. E.; *BBA-Protein Struct. M.* 2001, 1549, 123-136.
- (14) McDermott, G.; Prince, S. M.; Freer, A. A.; Hawthornthwaite-lawless, A. M.; Papiz, M. Z.; Cogdell, R. J.; Isaacs, N. W.; *Nature* 1995, 374, 517-521.
- (15) Egorova-Zachernyuk, T. A.; Hollander, J.; Fraser, N.; Gast, P.; Hoff, A. J.; Cogdell, R.; de Groot, H. J. M.; Baldus, M.; *J. Biomol. NMR* 2001, 19, 243-253.

- (16) van Gammeren, A. J.; Hulsbergen, F. B.; Hollander, J. G.; de Groot, H. J. M.; *J. Biomol. NMR* 2004, 30, 267-274.
- (17) Metz, G.; Wu, X. L.; Smith, S. O.; *J. Magn. Reson. Ser. A* 1994, 110, 219-227.
- (18) Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W. G.; Lansbury, P. T.; Griffin, R. G.; *J. Chem. Phys.* 1998, 108, 9463-9479.
- (19) Castellani, F.; van Rossum, B. J.; Diehl, A.; Rehbein, K.; Oschkinat, H.; *Biochemistry* 2003, 42, 11476-11483.
- (20) Baldus, M.; Petkova, A. T.; Herzfeld, J.; Griffin, R. G.; *Mol. Phys.* 1998, 95, 1197-1207.
- (21) Hediger, S.; Meier, B. H.; Ernst, R. R.; *Chem. Phys. Lett.* 1995, 240, 449-456.
- (22) Baldus, M.; Geurts, D. G.; Hediger, S.; Meier, B. H.; *J. Magn. Reson. Ser. A* 1996, 118, 140-144.
- (23) van Gammeren, A. J.; Hulsbergen, F. B.; Erkelens, C.; de Groot, H. J. M.; *J. Biol. Inorg. Chem.* 2004, 9, 109-117.
- (24) Alia; Matysik, J.; de Boer, I.; Gast, P.; van Gorkom, H. J.; de Groot, H. J. M.; *J. Biomol. NMR* 2004, 28, 157-164.
- (25) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 2003, 326, 1523-1538.
- (26) Alia; Matysik, J.; Soede-Huijbregts, C.; Baldus, M.; Raap, J.; Lugtenburg, J.; Gast, P.; van Gorkom, H. J.; Hoff, A. J.; de Groot, H. J. M.; *J. Am. Chem. Soc.* 2001, 123, 4803-4809.
- (27) Schulten, E. A. M.; Matysik, J.; Alia; Kiihne, S.; Raap, J.; Lugtenburg, J.; Gast, P.; Hoff, A. J.; de Groot, H. J. M.; *Biochemistry* 2002, 41, 8708-8717.
- (28) van Gammeren, A. J.; Buda, F.; Kiihne, S.; Hollander, J. G.; Hulsbergen, F. B.; Egorova-Zachernyuk, T. A.; Fraser, N. J.; Cogdell, R.; de Groot, H. J. M.; 2005, accepted for publication in the *J. Am. Chem. Soc.*

Chapter 4

Resolving the electronic structures of the B800 and B850 cofactors in the LH2 complex^{*}

4.1 Introduction

Bacteriochlorophylls (BChl) and carotenoids are the main light-absorbing cofactors in purple bacteria. They are non-covalently bound to two types of integral membrane proteins, the photosynthetic reaction centers and light-harvesting or antenna complexes.¹ The antenna cofactors absorb photons, yielding an excitation that is transferred via excitonic coupling and Förster energy transfer from the peripheral light-harvesting complexes to the reaction center where the charge separation takes place.²

Bacterial photosynthetic protein complexes can be isolated in relatively large quantities. They can be stabilized with a small amount of detergent. In this study is focused on resolving the electronic structure of the BChls in the bacterial photosynthetic light-harvesting 2 (LH2) complex from *Rhodospseudomonas (R.) acidophila* strain 10050 by Magic Angle Spinning (MAS) NMR. Detailed structural

^{*} The contents of this chapter have been published in the *Journal of the American Chemical Society*, 2005.

data have been obtained by X-ray crystallography for two bacterial photosynthetic LH2 complexes, from *R. acidophila* and from *Rhodospirillum rubrum*.^{3,4} The LH2 intrinsic membrane protein of *R. acidophila* is a circular aggregate complex containing nine identical monomeric units that consist of two membrane spanning helices, the α -subunit with 53 residues and the β -subunit with 41 residues. Each α and β -subunit binds three BChl *a* cofactors.³ One of the BChls connected to the α -subunit participates in a ring of nine BChls, which shows an absorption maximum at 800 nm that corresponds with an absorption energy of 12500 cm⁻¹. The remaining two BChls, one connected to the α -subunit and one connected to the β -subunit, participate in a concentric ring of 18 BChls, which has an absorption maximum at 860 nm corresponding with an absorption energy of 11627 cm⁻¹. Throughout the thesis is referred to these cofactors as the B800, α B850 and β B850 cofactors. The α B850 and β B850 in every monomeric unit form a pair of partly overlapping BChls that is sandwiched between the α - and β -subunits. Each pair also overlaps with the pair of the adjacent monomeric unit, forming a ring with continuous overlap between the 18 B850 cofactors. The nine B800 are positioned between the transmembrane helices of the β -subunits of each LH2 unit and form a ring without mutual overlap. The maximum absorption wavelengths of the B800 and B850 in the LH2 complex from *R. acidophila* are the Q_y-absorption bands. This refers to the Q_y dipole moment of the BChls, which can be physically mapped in the plane surface on the long axis of the π -bonding system in the ring system of the BChl cofactors. The Q_y-bands of the B800 and the B850 in the LH2 complex differ substantially from the spectrum of monomeric BChl *a* complexes in acetone : methanol 7:3 solution at 773 nm, corresponding with an absorption energy of 12937 cm⁻¹.⁵ They are 30 and 85 nm red shifted, respectively, while the Q_x absorption band, which represents the Q_x transition dipole moment that is coincident with the short axis of the π -bonding system, is hardly affected by the interaction with the protein. The Q_x band shows an absorption band at 590 nm for both monomeric BChl *a*, B800 and B850 complexes. It is known that protein-cofactor interactions and cofactor-cofactor interactions in the LH2 both contribute to the shift of the Q_y absorption band and hardly affects the Q_x absorption band.⁶

MAS NMR spectroscopy is able to specify the electronic structure of the BChls in the photoactive complexes. For instance, recently MAS NMR and photo-CIDNP were

combined and performed on photosynthetic reaction centers of *Rb. sphaeroides*, containing biosynthetic specifically ^{13}C enriched BChls in the special pair, and it was found that the electronic structure of the BChls in the reaction center is asymmetric.⁷ Recently, MAS NMR has also been used for a chemical shift assignment of the partly and fully labeled reconstituted retinylidene chromophore in rhodopsin and has revealed the electronic structure in the ground state of the native form.^{8,9} Here MAS dipolar correlation NMR spectroscopy is used for the selective assignment of the responses of BChls in the uniformly labeled LH2 protein complex isolated from *R. acidophila* strain 10050 and to resolve how the cofactor-cofactor interactions and the protein environment tune the electronic structures of these cofactors. The electronic differences for B800 and B850 in their natural environment have been resolved by comparing the solid-state isotropic ^{13}C chemical shifts ($\sigma^s_{B800}, \sigma^s_{B850}$) with the shifts for monomeric BChls dissolved in acetone- d_6 (σ^l). Density functional theory (DFT) calculations were performed to estimate the ring current effects for the B850 (σ^R). After correction for the ring current shifts, the molecular electronic ground state of the B850 could be characterized with atomic selectivity and global interactions with the LH2 protein matrix are resolved.

4.2 Materials and methods

Uniformly [^{13}C , ^{15}N] enriched LH2 complexes were obtained by growing *R. acidophila* strain 10050 anaerobically in light at 30 °C on a defined medium.¹⁰ The amounts of [U- ^{13}C , ^{15}N] labeled algae hydrolysate (Cambridge Isotopes Laboratories, Andover, MA, USA) and $^{13}\text{C}_4$ -labeled succinic acid were adjusted to 1.5 and 2.0 g/l, respectively, to optimize the cell growth. The $^{13}\text{C}_4$ -enriched succinic acid was prepared by a multi-step synthesis, starting with the synthesis of fumaric acid from $^{13}\text{C}_2$ -labeled acetic acid (CIL, Andover, MA, USA).¹¹ In the final step, fumaric acid was converted to succinic acid in a reduction reaction with H_2 and Pd-C as catalyst. After growing for 8 days, cells were harvested and the LH2 complexes were isolated.¹² The sample was concentrated to a volume of about 0.10 ml using a Centricon 100-kDa filter. The concentrated sample was transferred into a 4.0-mm CRAMPS rotor, containing approximately 10 mg of protein. The U- $^{13}\text{C},^{15}\text{N}$ -B800

LH2 complex was prepared by selective removal of the B800 molecule from the unlabeled LH2 complex, where the B800 binding site was reconstituted with U- ^{13}C , ^{15}N]-BChl α , using the reconstitution method described elsewhere.¹³

Radio frequency-driven dipolar recoupling (RFDR) and proton driven spin diffusion (PDS) techniques were used to record ^{13}C - ^{13}C correlation data with a Bruker AV-750 spectrometer using a double channel CP-MAS probe head. The proton $\pi/2$ pulse was set to 3.1 μs , corresponding with a nutation frequency of 80.6 kHz, while a ^{13}C B₁ field strength of 50 kHz with a cross polarization time of 2.0 ms was used during a 100-50% ramped CP sequence.¹⁴ 2D ^{13}C - ^{13}C RFDR spectra were acquired at a radio frequency of 188.6 MHz by using a pathway selective phase cycling method.¹⁵ The protons were decoupled by use of continuous wave (CW) decoupling of 80.6 kHz during the mixing time and by two-pulse phase modulation (TPPM) decoupling during t_1 and t_2 .¹⁶ Rotor-synchronized π -pulses with a length of 8.4 μs were applied during the mixing time. For optimal signal intensity in the aromatic region, the carrier frequency was set at 130 ppm on the chemical shift scale. For the PDS NMR experiment the proton decoupling was switched off during the spin diffusion mixing period to obtain ^1H mediated transfer of ^{13}C polarization along the molecular carbon network.

Samples were cooled to 223 K to immobilize the sample while maintaining good spectral resolution and the MAS spinning frequency $\omega_{\text{R}}/2\pi$ was 13 kHz. Additional PDS and RFDR NMR spectra were collected with spinning frequencies of 10.2 kHz to resolve correlations located at the spinning sidebands. Mixing times of 2.5 ms (RFDR) and 15 ms (PDS) were long enough to observe the correlations between directly bonded carbons as well as relayed correlations between more remote positions. To resolve all resonances, both 2D RFDR and PDS ^{13}C - ^{13}C correlation NMR spectra of the U- ^{13}C , ^{15}N] LH2 and the U- ^{13}C , ^{15}N]-B800 LH2 were recorded. The $^{13}\text{COOH}$ resonance of U- ^{13}C , ^{15}N]-tyrosine-HCl at 172.1 ppm was used as an external chemical shift reference to determine the isotropic chemical shifts (σ) of the B800/B850 carbons.

To calculate the local magnetic shielding due to the local magnetic field induced by the ring current effect in the BChls (σ^{R}), the nucleus independent chemical shift (NICS) calculations with the DFT method based on Becke's¹⁷ and Lee-Young-Parr's¹⁸ gradient-corrected correlation functional (BLYP) was used. Calculations were

performed at the BLYP/6-31G(d,p) level using the GIAO method as implemented in the Gaussian 98 package.¹⁹ The geometry of the B850 cofactors was taken from the X-ray without structure relaxation. To simplify the calculations, the phytyl tails connected to C17 were substituted by methyl groups. The σ^R_{B850} were calculated for each carbon which is in the shielding area of an adjacent overlapping B850. The coordinates of the carbons affected by the ring current were used as points in space and the shielding effect induced by the adjacent B850 was calculated at these positions. Four calculations were performed: two for the overlap between rings A, and two for the overlap between rings C (figure 4.2). The macrocycle of α B850 was used to calculate the shielding on carbons from the interacting β B850. In a next calculation, the macrocycle of the β B850 was used to calculate the shielding on carbons from the interacting α B850. This leads to independent calculations of the ring current effect without contributions from the π - π overlap and other electronic effects.

Electronic changes in B800 and B850 cofactors induced by the molecular environment were identified by measuring isotropic ^{13}C chemical shift differences ($\Delta\sigma$) between monomeric BChl *a* and B800/B850, defined as $\Delta\sigma = \sigma^s - \sigma^l$. For B800, the $\Delta\sigma_{B800}$ represents an estimate of the shift changes due to the cofactor-protein interactions. The mutual distances between the B800 are too large to induce significant ring current effects.²⁰ The $\Delta\sigma_{B850}$ for carbons in the B850 are corrected for the ring current contribution to the chemical shift σ^R_{B850} from the adjacent B850. This yields an estimate of the chemical shift change $\Delta\tilde{\sigma}_{B850}$ due to the interactions with the protein matrix, according to

$$\Delta\tilde{\sigma}_{B850} = \Delta\sigma_{B850} - \sigma^R_{B850} \quad (4.1)$$

4.3 Results

The high-resolution 2D RFDR ^{13}C - ^{13}C NMR spectrum of the U- $[^{13}\text{C}, ^{15}\text{N}]$ LH2 complex, shown in figure 4.1 (p. 65), can be subdivided into two parts. In the upper part of the figure, between 0 and 75 ppm, the correlations from the aliphatic protein response of the U- $[^{13}\text{C}, ^{15}\text{N}]$ LH2 complex are observed. The lower part, between 100

and 200 ppm, shows the correlation signals involving aromatic moieties. Most correlation signals from the BChl *a* cofactors in the LH2 complex can be resolved from the spinning sidebands by using a spinning frequency of 13 kHz. The various ^{13}C correlation signals for the aromatic ^{13}C carbons from the BChl *a* cofactors are labeled with a number that corresponds with the carbon positions in the macrocycle according to the IUPAC numbering shown in figure 4.2.

Aromatic carbons can be assigned from their nearest-neighbor correlations. In figure 4.1, the C3-C3¹ correlations coincide with the spinning side bands and they can be resolved in a dataset recorded with a spinning frequency of 10.2 kHz. The aliphatic C2¹, C7, C8, C12, C13², C17 and C18 can be assigned via correlations with adjacent aromatic carbons. In addition, the aliphatic carbons C7¹, C8¹, C8², C17¹, C17² and C18¹ are assigned from the long-range correlations to aromatic carbons. C18¹ and C17¹ are observed weakly in the upper part of the RFDR NMR spectrum, while correlations between C7¹, C8¹, C8² and C17¹ could be assigned from the crowded aliphatic region contour region in the PDSD NMR spectrum (data not shown). Due to relatively high local density of ^1H nuclei, which mediate the magnetization transfer in the PDSD technique, strong correlation signals of the aliphatic carbons are observed in the PDSD NMR spectrum that complement the RFDR dataset.

Finally there is strong overlap between the C17²-C17³ correlation signals and the broad correlation response from the carbonyl carbons between ~170 and ~180 ppm with the aliphatic C α and C β carbons between 0 and 75 ppm. This makes an assignment of the C17³ difficult. In contrast, the cofactor signals involving the C3¹, C13¹, C13³ have a good dispersion and are also well resolved by the long-range correlations due to the relayed transfer along the ^{13}C molecular framework.

Figure 4.1 (p. 65) Contour plot of a high-resolution MAS 2D ^{13}C - ^{13}C RFDR solid-state NMR spectrum of the $[U\text{-}^{13}\text{C},^{15}\text{N}]$ LH2 complex isolated from *R. acidophila* strain 10050. The data were obtained at 223 K with a spinning frequency of 13 kHz and a RFDR mixing time of 2.5 ms. The chemical shift correlation networks of the B800, the α B850 and the β B850 are indicated using the IUPAC numbering shown in figure 4.2. α B850 and β B850 are indicated by α and β , respectively.

For each BChl a cofactor species, a distinct network of nearest-neighbor carbon correlations can be identified by analyzing the correlations of the ring A carbons with carbons further in the molecular network by relayed transfer. With the help of another dataset collected from the $[^{13}\text{C},^{15}\text{N}]$ -B800 LH2 sample, the correlation network from the B800 can be assigned unambiguously. The two remaining datasets belong to the B850. The shift assignments are summarized in table 4.1.

The σ^R calculated with DFT for nuclei in the pyrrole rings A and C of the B850 cofactors are summarized in table 4.1. The σ^R 's are largest for C2, C2¹, C3, C3¹ and C3² from ring A and for positions C12, C12¹, C13 and C13¹ from ring C. The pyrrole rings A and C are within 3.9 Å and 3.6 Å of the macrocycle of an adjacent cofactor, respectively, and experience the largest local magnetic field effects induced by the ring current.³ A large difference between $\sigma^R_{\alpha\text{B850}}$ and $\sigma^R_{\beta\text{B850}}$ was calculated for carbons C2, C2¹, C3, C3¹ and C3² from ring A. The chemical shift differences between the α B850 and β B850 are small. The DFT calculations on the B850 cofactors indicate that the α B850 experiences a stronger shielding effect than the β B850. Hence the shift correlation set that is most shielded has been assigned to the α B850, while the other set is attributed to the β B850. Outside the overlap region the responses of the α B850 and β B850 are very similar, which shows that the electronic differences between α B850 and β B850 cofactors are small.

The shifts associated with electronic changes in the B800 relative to the monomeric BChl a in solution ($\Delta\sigma_{\text{B800}}$) can be determined directly from the observed $\Delta\sigma_{\text{B800}}$. The mutual distance of 21.2 Å between B800 cofactors is too large to exert a significant shielding effect on nearby B800 cofactors.²⁰

After applying the corrections for the ring current according to eq. 4.1 the shift differences between the B800, α B850 and β B850 are minimal. The $\Delta\sigma_{\text{B850}}$ of the ^{13}C in the bacteriochlorin ring are summarized in table 4.1, while the $\Delta\sigma_{\text{B850}}$ with $|\Delta\sigma_{\text{B850}}| \geq 1.0$ ppm are represented by proportional squares and circles in figure

4.2A. The solid and dashed lines in the figure represent the electronic effects on the α B850 and β B850, respectively. C17, C17¹ and C19 were not specifically assigned to α B850 or β B850. For these positions, the $\Delta\delta_{B850}$ are of the same order for the α B850 and the β B850 and are indicated in figure 4.2A with thick solid contours.

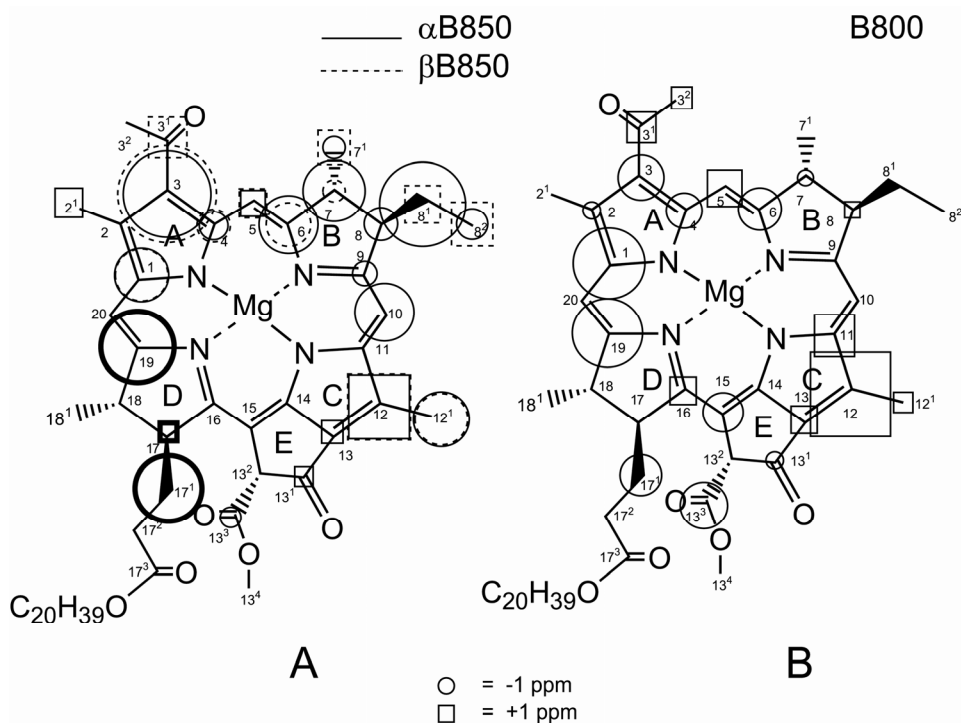


Figure 4.2 Estimated nuclear shift changes ($|\Delta\delta| \geq 1.0$ ppm), represented by circles (shielding effects) and squares (deshielding effects) for the α B850 and the β B850 (4.2A) and the B800 (4.2B) cofactors in the LH2 complex. The area of the mark is proportional to the magnitude of $\Delta\delta$.

4.4 Discussion

The changes of the chemical shifts with respect to monomeric BChl *a* in acetone-*d*₆ can be used to assess the electronic perturbations in the cofactors, which contributes to the red shifted Q_y bands of the B800 and the B850 cofactors

in their natural environment. In general, from the $\Delta\tilde{\sigma}_{B850}$ represented in figures 4.2A and 4.2B, it is evident that the electronic polarization effects in all three cofactor species are of the same order. The differences could arise from various mechanisms, such as structural differences or different protein-cofactor interactions. X-ray data shows that the β B850 is bent, while α B850 is planar. This may explain why minor shift differences between the two species in the NMR correlation spectrum are observed. The NMR spectra show that this structural difference provokes very small electronic differences between the ground states of the two B850 species. This contrasts with recent model studies by Fajer et. al. and Senger et. al.^{21,22} Their studies demonstrated that in principle structural distortions can have a pronounced effect on the electronic structure, although the distortions for their molecules were much larger than for the B850 macrocycles in the LH2 complex.

The strongest negative $\Delta\tilde{\sigma}$ are detected for the π -carbons located in and close to ring A, while pronounced positive $\Delta\tilde{\sigma}$ are detected for carbons in ring C. The negative $\Delta\tilde{\sigma}$ are distributed over aromatic carbons located between C6 and C19 around ring A, while the positive $\Delta\tilde{\sigma}$ distribution is restricted to predominantly C11, C12 and C13 in ring C. The chemical shift changes indicate a stabilization of partly negative charge density (δ^-) at and around ring A and polarized positive charge (δ^+) at ring C, corresponding with an induced electronic dipole moment parallel to the long axis of the π -bonding system and to the Q_y transition dipole moment. The difference between the negative induced charge at the vicinity of ring A and the positive induced charge at ring C induces approximately 20-25 ppm cumulative shift effects for both cofactors species, representing comparable dipole strengths along the long axis of the bacteriochlorin ring in the ground states. The dipole of the α B850 is slightly different compared to the others, due to the relative large negative $\Delta\tilde{\sigma}$ observed at C9 and C10.

The electronic effects are distributed over a relatively large part of the molecular structure of the B800 and B850 macrocycles. This implies that local interactions, like the hydrogen bond interaction of the C3-acetyl group and the axial coordination of the Mg can not explain the color shifts in the LH2. The NMR data provide evidence for global mutual polarization effects between the cofactors and their protein environment. In this picture, the polarity of the environment affects the

dielectric properties of the LH2 protein interior, leading to a collective effect on the B800 and B850 electronic structure. Local interactions can still be essential structural elements provoking changes of the protein functional characteristics upon removal of the hydrogen bonding groups as shown by mutation studies.²³⁻²⁸ A clear example is provided for the B800/B820 or LH3 complex, which is a spectroscopic variant of the LH2.²⁸ In the active binding site of the cofactors of this complex, a specific amino acid holds the C3-acetyl groups in an out-of-plane orientation, thereby changing the properties of the cofactors. DFT calculations performed on α B850 indicate a significant blue-shift for the BChls with the C3-acetyl group perpendicular to the plane of the macrocycle with regard to the C3-acetyl in a coplanar orientation. This is in line with previous theoretical calculations.²⁹

Mutation studies on the protein environment that surrounds the cofactors of the reaction center in *Rb. sphaeroides* have shown that the dielectrics of environment of the cofactors is affected by charge separation in a non-linear response mechanism due to synergistic contributions of various chemical functional polar groups, leading to changes of the absorption wavelengths of the cofactors.³⁰ This has been confirmed by chemical modeling, which includes the dielectrics of the protein environment of the photosynthetic reaction centre.³¹

The narrow NMR signals of the bacteriochlorin ring carbons in the 2D RFDR NMR spectrum show that the LH2 complexes in the sample are well ordered and all LH2 units of the nonameric aggregate complex are from the NMR point of view spectroscopically identical. When viewed on a fast time scale in the excited state, disorder can be observed, as has been demonstrated by time resolved optical studies.³²⁻³⁵ The NMR study shows narrow resonances. This shows that the ground state structure on the much longer NMR time-scale is on average very regular, in line with X-ray studies. If there are differences between the site energies of the individual BChl cofactors, the effect on the electronic structure should be small compared to the difference between the α B850 and β B850, which is clearly resolved in the NMR spectra. Of the long phytol chains, only the p1, p2, p3 and p17 carbons are detected, which suggests that the tails may be disordered in the frozen sample.

Since the Q_y band of B800 has a red shift of about 30 nm (437 cm^{-1}) and the induced electronic effects and corresponding electronic dipoles in the B800 and

B850 are similar, it is difficult to attribute the additional red shift of 60 nm (873 cm⁻¹) for the B850 cofactors to electronic changes of the ground state. Structure-based calculations performed on LH2 that use a point monopole approximation to represent B800 and B850 Q_y dipole transition moments, indicate that the coupling between the 18 B850 rings in LH2 yields a 40 nm red shift compared to the Q_y band of the ring of nine B800, assuming that the site energies are identical for both cofactor species.³⁶ This corresponds to a Q_y band at 840 nm for the B850 cofactors. When the $\Delta\tilde{\sigma}$ on the B800 are associated to the ~30 nm red shift from monomeric BChl *a* in solution to the protein environment, it is virtually impossible that the small $\Delta\tilde{\sigma}$ differences between the B800 and B850 represent the B800-B850 absorption difference. Hence the large B800-B850 absorption difference should be attributed predominantly to excitonic coupling between the transition dipole moments of the B850 cofactors.^{34,35}

In addition to the changes of the electric dipole moment of the rings, the cumulative $\Delta\tilde{\sigma} < 0$ imply a negative charge on the bacteriochlorin rings. This is in line with recent studies that have shown that the imidazole moieties of the Mg-coordinating histidines in the LH2 protein are positively charged, since the positive charge at the imidazole is compensated and stabilized by negative charge at the BChl.^{37,38} The cumulative $\Delta\tilde{\sigma}$ are -8.1, -14.4 and -9.4 ppm for B800, α B850 and β B850, respectively. A detailed analysis of the electronic effects shows that the balances between the positive and negative induced charge effects in B800 and B850 are slightly different. Comparing the positive $\Delta\tilde{\sigma}$ effects in the vicinity of ring C clearly suggests that this area in the B800 is more positively charged than the corresponding area of the α B850 and β B850. The cumulative $\Delta\tilde{\sigma} < 0$ effects on the aromatic carbons in the vicinity of ring A are similar for all three cofactor species. Pronounced $\Delta\tilde{\sigma}$ differences between the negative $\Delta\tilde{\sigma}$ distributions in the B800 and B850 cofactors are observed at C1 and C3. In the B800 complexes, a strong $\Delta\tilde{\sigma}$ effect is observed for C1 and a moderate $\Delta\tilde{\sigma}$ effect is detected for C3, while the reverse effect for those carbons is observed in the B850. The C3-acetyl groups of the B850 are twisted compared to the C3-acetyl group of B800 and may contribute to the chemical shift differences.

Table 4.1 Summary of the isotropic chemical shift (σ), the isotropic chemical shift of BChl *a* monomers in acetone (σ^A), the isotropic chemical shift difference between monomeric BChl in acetone and cofactors in the LH2 complex ($\Delta\sigma$), the ring current shift calculated with DFT (σ^R), and the chemical shift change due to electronic effects ($\Delta\sigma^e$). ^a Correlation peaks are not observed.

Carbon	σ_{B800}	$\sigma_{\alpha B850}$	$\sigma_{\beta B850}$	σ^I	$\Delta\sigma_{B800}$	$\Delta\sigma_{\alpha B850}$	$\Delta\sigma_{\beta B850}$	$\sigma^R_{B850}(\alpha,\beta)$	$\Delta\sigma^e_{\alpha B850}$	$\Delta\sigma^e_{\beta B850}$
1	146.2	147.1	147.8	151.2	-5.0	-4.1	-3.4	-0.75, -0.14	-3.4	-3.3
2	141.0	141.1	141.3	142.1	-1.1	-1.0	-0.8	-1.93, -0.86	0.9	0.1
2 ¹	13.3	12.3	12.2	13.5	-0.2	-1.2	-1.3	-3.10, -1.41	1.9	0.1
3	134.4	130.3	130.7	137.6	-3.2	-7.3	-6.9	-1.71, -0.74	-5.6	-6.2
3 ¹	201.4	197.9	201.1	199.3	2.1	-1.4	1.8	-1.72, -0.72	0.3	2.5
3 ²	34.4	29.2	31.9	32.9	1.5	-3.7	-1.0	-2.81, -1.32	-0.9	0.3
4	147.7	147.7	147.9	150.2	-2.5	-2.5	-2.3	-0.70, -0.19	-1.8	-2.1
5	102.1	100.8	101.5	99.6	2.5	1.2	1.9	-0.40, -0.14	1.6	1.8
6	165.9	165.2	166.1	168.9	-3.0	-3.7	-2.8		-3.7	-2.8
7	47.2	44.3	46.8	48.2	-1.0	-3.9	-1.4		-3.9	-1.4
7 ¹	^a	22.1	25.3	23.1		-1.0	2.2		-1.0	2.2
8	56.6	53.5	56.1	55.6	1.0	-2.1	0.5		-2.1	0.5
8 ¹	^a	25.4	32.3	30.8		-5.4	2.3		-5.4	2.3
8 ²	^a	9.1	13.2	10.5		-1.4	2.7		-1.4	2.7
9	157.8	157.4	159.2	158.5	-0.7	-1.1	0.7		-1.1	0.7
10	102.6	99.0	102.5	102.4	0.2	-3.4	0.1	0.27, 0.05	-3.7	0.1
11	152.4	150.5	150.1	149.5	2.9	1.0	0.6	0.10, -0.21	0.9	0.8
12	129.7	127.3	126.9	123.9	5.8	3.4	3.0	-0.56, -1.05	4.0	4.1
12 ¹	13.3	7.2	6.6	11.9	1.4	-4.7	-5.3	-1.25, -2.04	-3.5	-3.3
13	132.4	131.4	129.9	130.5	1.9	0.9	-0.6	-0.50, -0.78	1.4	0.2
13 ¹	187.9	189.7	189.1	189.0	-1.1	0.7	0.1	-0.46, -0.55	1.2	0.7
13 ²	65.3	65.2	64.7	65.7	-0.4	-0.5	-1.0	-0.06, -0.09	-0.4	-0.9
13 ³	168.3	170.3	172.0	171.6	-3.3	-1.3	0.4		-1.3	0.4
14	160.1	160.0	161.0	160.8	-0.7	-0.8	0.2	-0.02, -0.22	-0.8	0.4
15	106.9	109.4	110.1	109.7	-2.8	-0.3	0.4	0.03, -0.07	-0.3	0.4
16	154.1	153.1	152.3	152.2	1.9	0.9	0.1		0.9	0.1
17	49.7	51.6, 51.7		50.5	-0.8		1.1, 1.2			
17 ¹	27.6	26.8, 26.3		30.5	-2.9		-3.7, -4.2			
17 ²	^a	^a		29.4						
17 ³	^a	^a		173.4						
18	49.1	49.8, 49.6		49.3	-0.2		0.5, 0.3			
18 ¹	24.2	24.2, 24.2		23.3	0.9		0.9, 0.9			
19	162.4	162.3, 162.7		167.3	-4.9		-5.0, -4.6			
20	96.9	97.2, 97.6		96.3	0.6		0.9, 1.3			

Finally, shift differences are observed for C7, C7¹, C8, C8¹, C8², C13¹, C17, and C17¹ of α B850 and β B850. It is believed that these effects, in particular for C8, C8¹, and C8², reflect conformational changes of the aliphatic side-chains of the macrocycle enforced by interactions with the protein. The $\Delta\sigma$ of -2.7 and -3.3 ppm for the carbons C12¹ of the B850 cofactors are anomalously large. X-ray shows that that both B850 C12¹ carbons are located above adjacent B850 macrocycle. However, this leads to somewhat smaller σ^R of -1.25 and -2.04 ppm (table 4.1). The imidazole rings of the coordinating histidines also induce a local magnetic field, which can produce a small additional ring current effect of about -0.5 ppm on the C12¹ carbons. However, this does not fully explain the total $\Delta\sigma$ for C12¹.³⁹

4.5 Conclusion

Isotropic chemical shift assignments have been obtained for the B800 and B850 cofactors in the LH2 transmembrane protein complex to resolve their molecular electronic structures with atomic selectivity. Three resonance sets from the B800 and two B850 cofactors were extracted selectively from the spectrum of the uniformly [¹³C, ¹⁵N]-labeled LH2 complex isolated from *R. acidophila* strain 10050. The NMR dataset of B800 was selectively and unambiguously assigned with the help of a reconstituted U-[¹³C, ¹⁵N]-B800 LH2 complex.

Isotropic chemical shift datasets obtained from the solid-state NMR spectra indicate different chemical environments for each cofactor species. After correcting the isotropic shifts for the ring currents in the B850 cofactors, the differences between the three cofactors are small and reveal electronic similarity between B800, α B850 and β B850 cofactors. All three species show a similar polarization of the long axis of the bacteriochlorin ring parallel to the Q_y transition dipole moment. This contrasts with the special pair in the reaction centers of *Rb. sphaeroides*, where convincing evidence was provided for an asymmetric electronic structure of the BChls in the special pair in the reaction center.⁷ A similar electronic dipole for both the B800 and B850 species is detected by NMR, which leads to the conclusion that only a part of the red shift, about 30 nm, can be attributed to the protein environment. The large additional red shift of the B850 absorption band with

respect to B800 is mainly due to overlap between the B850 bacteriochlorin rings in the LH2 protein complex.

The electronic charge polarization, which is mainly along the long axis of the π -bonding system, is due to a collective effect of the synergistic contributions of the polar moieties in the cofactor environment. Since shift effects are distributed over the bacteriochlorin ring, rather than localized on specific carbon positions, the induced polarization is attributed to mutual polarization between the cofactors and the protein in a nonlinear dielectric response mechanism. A global collective molecular electronic response rather than specific localized effects is in line with theoretical studies that point out the importance of the dielectrics for the protein interior.³¹

References

- (1) Cogdell, R. J.; Isaacs, N. W.; Freer, A. A.; Arrelano, J.; Howard, T. D.; Papiz, M. Z.; Hawthornthwaite-Lawless, A. M.; Prince, S.; *Prog. Biophys. Mol. Biol.* 1997, 68, 1-27.
- (2) Hoff, A. J.; Deisenhofer, J.; *Phys. Rep.* 1997, 287, 2-247.
- (3) Papiz, M. Z.; Prince, S. M.; Howard, T.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 2003, 326, 1523-1538.
- (4) Koepke, J.; Hu, X. C.; Muenke, C.; Schulten, K.; Michel, H.; *Structure* 1996, 4, 581-597.
- (5) Scheer, H.; In: *The Chlorophylls* CRC Press: Boca Raton FL, 1991.
- (6) Cogdell, R. J.; Howard, T. D.; Isaacs, N. W.; McLuskey, K.; Gardiner, A. T.; *Photosynth. Res.* 2002, 74, 135-141.
- (7) Schulten, E. A. M.; Matysik, J.; Alia; Kiihne, S.; Raap, J.; Lugtenburg, J.; Gast, P.; Hoff, A. J.; de Groot, H. J. M.; *Biochemistry* 2002, 41, 8708-8717.
- (8) Verhoeven, M. A.; Creemers, A. F. L.; Bovee-Geurts, P. H. M.; De Grip, W. J.; Lugtenburg, J.; de Groot, H. J. M.; *Biochemistry* 2001, 40, 3282-3288.
- (9) Creemers, A. F. L.; Kiihne, S.; Bovee-Geurts, P. H. M.; DeGrip, W. J.; Lugtenburg, J.; de Groot, H. J. M.; *Proc. Natl. Acad. Sci. U. S. A.* 2002, 99, 9101-9106.

- (10) Egorova-Zachernyuk, T. A.; Hollander, J.; Fraser, N.; Gast, P.; Hoff, A. J.; Cogdell, R.; de Groot, H. J. M.; Baldus, M.; *J. Biomol. NMR* 2001, 19, 243-253.
- (11) Heinen, W.; In: *Grafting of polyolefins and miscibility in copolymer mixtures*; Leiden University, Leiden, 1996.
- (12) Hawthornthwaite-Lawless, A. M.; Cogdell, R. J.; In: *The Chlorophylls*; Scheer, H., Ed.; CRC: Boca Raton, 1991; pp 494-528.
- (13) Fraser, N. J.; Dominy, P. J.; Ucker, B.; Simonin, I.; Scheer, H.; Cogdell, R. J.; *Biochemistry* 1999, 38, 9684-9692.
- (14) Metz, G.; Wu, X. L.; Smith, S. O.; *J. Magn. Reson. Ser. A* 1994, 110, 219-227.
- (15) Boender, G. J.; Vega, S.; de Groot, H. J. M.; *J. Chem. Phys.* 2000, 112, 1096-1106.
- (16) Bennett, A. E.; Rienstra, C. M.; Griffiths, J. M.; Zhen, W. G.; Lansbury, P. T.; Griffin, R. G.; *J. Chem. Phys.* 1998, 108, 9463-9479.
- (17) Becke A. D.; *J. Chem. Phys.* 1986, 84, 4524.
- (18) Lee, C. T.; Yang, W. T.; Parr, R. G.; *Phys. Rev. B* 1988, 37, 785-789.
- (19) Frisch M.J.; Trucks G.W.; Schlegel H.B.; Scuseria G.E.; Robb M.A.; Cheeseman J.R.; Zakrzewski V.G.; Montgomery Jr. J.A.; Stratmann R.E.; Burant J.C.; Dapprich S.; Millam J.M.; Daniels A.D.; Kudin K.N.; Strain M.C.; Farkas O.; Tomasi J.; Barone V.; Cossi M.; Cammi R.; Mennucci B.; Pomelli C.; Adamo C.; Clifford S.; Ochterski J.; Petersson G.A.; Ayala P.Y.; Cui Q.; Morokuma K.; Malick D.K.; Rabuck A.D.; Raghavachari K.; Foresman J.B.; Cioslowski J.; Ortiz J.V.; Stefanov J.B.; Liu G.; Liashenko A.; Piskorz P.; Komaromi I.; Gomperts R.; Martin R.L.; Fox D.J.; Keith T.; Al-Laham M.A.; Peng C.Y.; Nanayakkara A.; Gonzalez C.; Challacombe M.; Gill P.M.W.; Johnson B.; Chen W.; Wong M.W.; Andres J.L.; Gonzalez C.; Head-Gordon M.; Replogle E.S.; Pople J.A.; *Gaussian 98 revision A.5, Inc., Pittsburgh, PA*, 1998.
- (20) Freer, A.; Prince, S.; Sauer, K.; Papiz, M.; Hawthornthwaite-Lawless, A.; McDermott, G.; Cogdell, R.; Isaacs, N. W.; *Structure* 1996, 4, 449-462.
- (21) Fajer, J.; *J. Porphyr. Phthalocya.* 2000, 4, 382-385.

- (22) Senge, M. O.; Renner, M. W.; Kalisch, W. W.; Fajer, J.; *J. Chem. Soc. Dalton Trans.* 2000, 381-385.
- (23) Fowler, G. J. S.; Visschers, R. W.; Grief, G. G.; van Grondelle, R.; Hunter, C. N.; *Nature* 1992, 355, 848-850.
- (24) Fowler, G. J. S.; Sockalingum, G. D.; Robert, B.; Hunter, C. N.; *Biochem. J.* 1994, 299, 695-700.
- (25) Fowler, G. J. S.; Hess, S.; Pullerits, T.; Sundstrom, V.; Hunter, C. N.; *Biochemistry* 1997, 36, 11282-11291.
- (26) Gall, A.; Ridge, J. P.; Robert, B.; Cogdell, R. J.; Jones, M. R.; Fyfe, P. K.; *Photosynth. Res.* 1999, 59, 223-230.
- (27) Gall, A.; Fowler, G. J. S.; Hunter, C. N.; Robert, B.; *Biochemistry* 1997, 36, 16282-16287.
- (28) McLuskey, K.; Prince, S. M.; Cogdell, R. J.; Isaacs, N. W.; *Biochemistry* 2001, 40, 8783-8789.
- (29) Gudowska-Nowak, E.; Newton, M. D.; Fajer, J.; *J. Phys. Chemistry* 1990, 94, 5795-5801.
- (30) Moore, L. J.; Zhou, H. L.; Boxer, S. G.; *Biochemistry* 1999, 38, 11949-11960.
- (31) Sham, Y. Y.; Muegge, I.; Warshel, N.; *Biophys. J.* 1999, 76, A198-A198.
- (32) van Oijen, A. M.; Ketelaars, M.; Matsushita, M.; Kohler, J.; Aartsma, T. J.; Schmidt, J.; *Biophys. J.* 2001, 80, 151A-151A.
- (33) van Oijen, A. M.; Ketelaars, M.; Kohler, J.; Aartsma, T. J.; Schmidt, J.; *Chem. Phys.* 1999, 247, 53-60.
- (34) van Oijen, A. M.; Ketelaars, M.; Kohler, J.; Aartsma, T. J.; Schmidt, J.; *Science* 1999, 285, 400-402.
- (35) Georgakopoulou, S.; Frese, R. N.; Johnson, E.; Koolhaas, C.; Cogdell, R. J.; van Grondelle, R.; van der Zwan, G.; *Biophys. J.* 2002, 82, 2184-2197.
- (36) Sauer, K.; Cogdell, R. J.; Prince, S. M.; Freer, A.; Isaacs, N. W.; Scheer, H.; *Photochem. Photobiol.* 1996, 64, 564-576.
- (37) van Gammeren, A. J.; Hulsbergen, F. B.; Erkelens, C.; de Groot, H. J. M.; *J. Biol. Inorg. Chem.* 2004, 9, 109-117.

- (38) Alia; Matysik, J.; Soede-Huijbregts, C.; Baldus, M.; Raap, J.; Lugtenburg, J.; Gast, P.; van Gorkom, H. J.; Hoff, A. J.; de Groot, H. J. M.; *J. Am. Chem. Soc.* 2001, 123, 4803-4809.
- (39) Giessner-Prettre, C.; Pullman, B.; *J. Theor. Biol.* 1971, 31, 287-294.

Chapter 5

Chemical model systems of the BChl α -histidine complex*

5.1 Introduction

Chlorophyll (Chl) and bacteriochlorophyll (BChl) are the most abundant pigments of photosynthesis, which is the energetic basis of life on earth. They are mainly found in the protein complexes of photosynthetic organisms, and are functional elements in light harvesting, energy transfer and electron transfer.¹ They are usually encapsulated in the light-harvesting (LH) and photosynthetic reaction center complexes, which are the principal elements of the photosynthetic units of green plants and bacteria.²

(B)Chl contains an Mg ion in the center of the porphyrin macrocycle that is coordinated by four pyrrole nitrogens (figure 5.1). Two axial positions are available for the coordination of ligands to the Mg. From vibrational spectra and X-ray crystallographic analyses it was found that the Mg ions of most (B)Chls in photosynthetic proteins are five-coordinated by the N τ -atom from a histidyl residue

* The contents of this chapter have been published in the *Journal of Biological Inorganic Chemistry*, 2004, 109-117.

in the protein. Due to this coordination, the Mg is slightly out of the equatorial plane.^{3,4} It has been reported that the Mg out of plane distance of the five-coordinated [Mg-OEP-pyridine] complex is 0.72-0.81 Å and since the Mg ion of the five-coordinated structure is somewhat out of plane, a second axial ligand coordination at the opposite side is thought to be less favorable.⁵ Six-coordinated species are only formed in concentrated ligand solutions and in polar donor solvents that form solvate complexes and compete with the ligand for coordination to the Mg. It has been proposed that polar histidyl residues can be involved in stabilization of the charge separation process in photosynthesis.⁶ This is supported by model studies that argued that charge-charge interactions between polar groups in the reaction center of *Rb. sphaeroides* can give rise to a non-linear increase of the dielectric constant of the protein interior and stabilize a charge state.⁷

A model study in which NMR spectra were recorded on codissolved Im and (B)Chl a in THF, has proposed a negatively charged [(B)Chl a-Im] complex due to the deprotonation of the Im NH1 hydrogen.⁸ The coordination behavior of imidazole and 1-methyl-imidazole (Im and 1-MeIm) to the Mg of a series of Mg-Por systems, including (B)Chl a, is here investigated by model studies in an attempt to mimic the neutral Mg-histidyl coordination which is present in photosynthetic proteins. In a more general perspective, the information from these models is potentially also of interest for other metal-containing complexes like the heme proteins and vitamin B12.

¹H and ¹³C NMR spectroscopy are used to monitor ¹H and ¹³C coordination shifts ($\Delta\delta$) for different [ligand] : [Mg-Por] molar ratios (r_m). The $\Delta\delta$ is the difference between the isotropic shifts of the free and the coordinated forms of the ligand and is due to the effect of the ring currents in the macrocycle of the porphyrin systems. In addition to ¹H NMR spectroscopy, ¹⁵N NMR and ¹³C NMR can be applied to provide insight into the structure and electronic properties of the axially coordinated ligand. In particular ¹⁵N-NMR experiments are of interest, since the Mg and the Im N-donor atom form a coordination bond, and the δ_N is very sensitive to the electronic configuration. In this work, ¹³C and ¹⁵N-NMR experiments are performed, using [1,3-¹⁵N, 2-¹³C]-Im and [1,3-¹⁵N]-Im, [1,2-¹³C]-1-MeIm. ¹⁵N labeling at both positions is relatively straightforward, while ¹³C nuclei can be incorporated conveniently at positions 1 and 2.⁹

The pyrrole-type (N1) and pyridine-type (N3) nitrogens, present in Im and 1-MeIm, each have very characteristic chemical shift ranges in the ^{15}N NMR spectra. The pyrrole-type signal is shifted by about 100 ppm from the resonance of the pyridine-type.¹⁰ For Im in protic solvents or in concentrated solutions, the proton exchanges very rapidly in a tautomeric equilibrium, yielding an average chemical shift $\bar{\delta}^{-15}\text{N}$. In contrast, 1-MeIm shows two different ^{15}N resonances, due to the presence of the methyl group. It has been reported that substitution of the hydrogen at the pyrrole-type nitrogen by a methyl group has only a marginal effect on the isotropic shift of the pyrrole-type nitrogen.^{11,12}

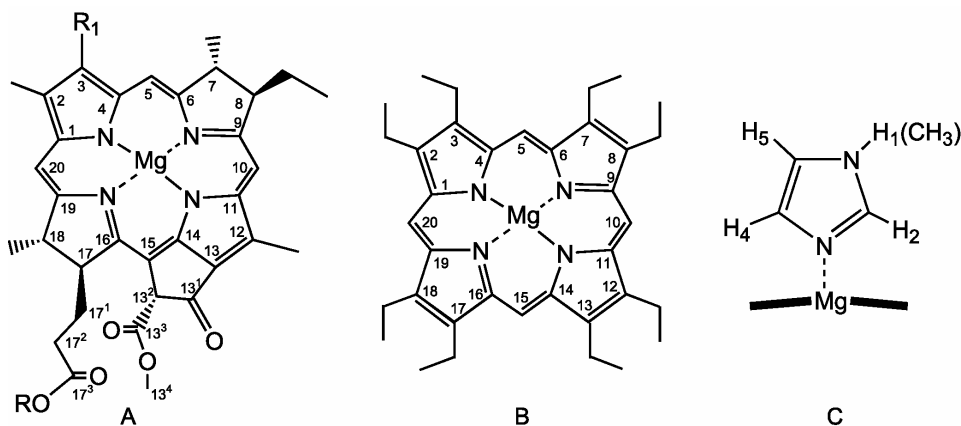


Figure 5.1 A) The structure of BChl *a* : $R = \text{phytyl}$, $R_1 = \text{COCH}_3$ and Chl *a* : $R = \text{phytyl}$, $R_1 = \text{CH}=\text{CH}_2$. B) The structure of Mg(II)-OEP. C) A schematic figure of the axial coordination of a diazole ligand to the Mg(II)-porphyrin ring.

5.2 Experimental

5.2.1 Synthesis

Labeled starting materials, all with 99% isotope incorporation, were purchased at Cambridge Isotopes Laboratories (CIL, Andover, MA, USA) and used without further purification. For the synthesis of isotopically enriched Im and 1-MeIm a known procedure was followed to obtain 1.08 g (15.4 mmol) of [1,3- ^{15}N 2- ^{13}C]-Im, 1.00 g [1,3- ^{15}N]-Im (15.4 mmol) and 0.228 g (2.65 mmol) of [1,3- ^{15}N 1,2- ^{13}C]-1-MeIm.⁹ The

purity of Im obtained by two consecutive recrystallisations was about 99% and the material was used without further purifications. To obtain a completely dry ligand solution for NMR investigations, 25 mg of ligand was dissolved in 5.0 ml CDCl₃ and dried over activated molecular sieves 3Å in a dry argon atmosphere.

59.0 mg (0.11 mmol) of Mg-OEP was prepared, starting from 100 mg (0.19 mmol) metal-free OEP (Porphyrin Systems) and an excess of MgBr₂·OEt₂ (490 mg, 1.90 mmol) in CH₂Cl₂.¹³ The compound was washed with water. Subsequently, the Mg-OEP was eluted rapidly from an alumina column (10 x 1.5 cm) with acetone p.a. grade, giving 59.0 mg (0.11 mmol) of 97:3 % Mg-OEP/metal-free OEP.

The complex of Mg-OEP with one axial Im or 1-MeIm ligand was achieved by dissolving Mg-OEP with just sufficient warm *iso*-butanol. To this solution a slight excess of ligand solution was added, and some drops of triethyl formate were added for dehydration. After 1-2 days, the [Mg-OEP·1,3-¹⁵N, 1,2-¹³C-1-MeIm] complex appeared as small crystals in the solution. To obtain the crystalline [Mg-OEP·1,3-¹⁵N-Im] complex, an amount of petroleum ether 100-120 equal to the *iso*-butanol volume was added to the complex solution and the mixture was placed for 14 days at -35 °C. After crystallization the supernatant was decanted.

5.2.2 (B)Chl *a* isolation

A mixture of Chl *a* and Chl *b* was extracted from spinach leaves, using an efficient method described by *Omata et. al.*¹⁴ To separate Chl *a* from Chl *b*, a silica column (3.0 x 12 cm), was used. The fractions were slowly eluted with hexane/propanol (20:1). The first band eluted from the column contained pure Chl *a*; the second band contained pure Chl *b*. The solvent of the Chl *a* solution was evaporated and the residue was stored under argon at -35 °C until it was used.

BChl *a* was isolated from a *Rb. sphaeroides* R26 culture.¹⁵ Briefly, bacteria were grown anaerobically in a sterile synthetic medium at 30 °C and with a light intensity of 2700 lux. Each liter of medium contained: 0.75 g of algae hydrolysate, 1.00 g succinic acid, 2.6 ml 6.5 M NH₄OH, 20 ml of 1.0 M H₂KPO₄, 20 ml of 1.0 M HK₂PO₄, 0.60 g MgSO₄·2H₂O, 0.12 g NaCl, 0.050 g CaCl₂·2H₂O, 125 µl vitamin solution (1.0 g thiamine·HCl + 8.0 mg biotine in 20 ml 5% EtOH/H₂O), 10 mg nicotinic acid, 10 ml of a trace element stock solution containing 316 mg/l FeCl₂, 20 mg/l MnSO₄·H₂O, 10 mg/l H₃BO₃, 10 mg/l CuSO₄·5H₂O, 10 mg/l

Na₂MoO₄·2H₂O, 20 mg/l ZnCl₂, 200 mg/l CaCl₂·2H₂O and 40 mg/l MgSO₄·7H₂O. The pH of the medium was adjusted with concentrated HCl to 6.8.

After growing for 8 days, the cells were harvested and the absorption of the cell culture was measured at 863 nm (A^{863} : 3.1/cm). The cells were centrifuged (20 min at 15000 $\times g$). The cell pellet was suspended in 20 mM Tris-HCl buffer (pH = 8.5) and stored at -35°C. The cell pellet was thawed and a mixture of acetone/methanol (7:3) was added to extract the chromophores from the cells. The suspension was centrifuged to precipitate the cell fragments. The pigment containing supernatant was concentrated and purified with a cross-linked agarose column of 1.5 x 2.0 cm) (Sephacose CL-6B). First a thin green band was eluted from the column, using hexane/acetone 20:1 as eluents. Subsequently, the intense blue band, containing the BChl *a*, was eluted easily from the column with pure acetone. The solution was concentrated and the pure BChl *a* was stored under argon at -35 °C until use.

5.2.3 Sample preparations

Activated molecular sieves (3Å) were used to prepare anhydrous CDCl₃. To exclude oxygen and moist from air, the sample preparation was performed under anaerobic conditions by using dry argon gas and a Schlenck apparatus. Experiments with (B)Chl *a* were performed in a dark and cold room (5 °C).

For each ¹H and ¹³C NMR experiment, about 1-2 mg (B)Chl *a* was dissolved in 2 ml of 2-propanol for the azeotropic removal of traces of water and transferred into a two-necked round bottom flask. The solvent was evaporated and the concentrate was exposed to vacuum (0,03 mm Hg) at 40 °C for two hours. Subsequently, argon gas was led in and 500 μl of anhydrous CDCl₃ was added. A NMR tube was annealed with a burner and placed in a Schlenck apparatus. After cooling to ambient temperature under high vacuum, argon gas was led in and the prepared solution was transferred into the NMR tube. The NMR tube was covered with a rubber stopper to keep the sample in a dry argon atmosphere. For the preparation of a water-free sample of Mg-OEP, the Mg-OEP (~1-2 mg) was weighed in a NMR tube before the annealing procedure.

5.2.4 NMR and IR experiments

For ^1H and ^{13}C NMR experiments, various amounts of 0.70 mM Im or 0.60 mM 1-MeIm ligand solution, ranging from 10 to 450 μl , were injected into the NMR sample tube by using a micro syringe. After each injection of ligand solution, a ^1H and ^{13}C NMR spectrum was recorded to determine the $\Delta\delta_{\text{H}} < 0$ and $\Delta\delta_{\text{C}} < 0$ of the ligand nuclei, relative to the corresponding shifts for the ligand molecules in solution. The $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{C}}$ of the ligand hydrogens were measured for various r_m , starting at the lowest ligand concentration. The r_m was determined from the integrals of the resonances in the ^1H NMR spectra.

All spectra were recorded at 298 K with a Bruker DPX spectrometer operating at 9.4 T. The δ_{H} of CHCl_3 at 7.26 ppm was used for internal calibration of the ^1H shift scale, while the response of CDCl_3 at 77.0 ppm was used for the calibration in the ^{13}C NMR spectra. For ^{15}N NMR experiments, the crystalline $[\text{Mg-OEP}\cdot 1,3\text{-}^{15}\text{N}, 1,2\text{-}^{13}\text{C}\text{-}1\text{-MeIm}]$ and the $[\text{Mg-OEP}\cdot 1,3\text{-}^{15}\text{N-Im}]$ complex were dissolved in anhydrous CDCl_3 . A relaxation delay of 10 s was used after each $\pi/4$ pulse. The δ_{N} is referenced relative to the response of liquid NH_3 with $\delta_{\text{N}} = 0.0$ ppm. $^1\text{H}\text{-}^{15}\text{N}$ correlations were determined by using the heteronuclear multiple bond correlation (HMBC) solution NMR technique. A ^{15}N solid-state MAS NMR spectrum from the crystalline $[\text{Mg-OEP}\cdot 1,3\text{-}^{15}\text{N}, 1,2\text{-}^{13}\text{C}\text{-}1\text{-MeIm}]$ complex prepared in a 4.0 mm CPMAS rotor was recorded, using a DMX-400 Bruker NMR spectrometer. The $\Delta\delta_{\text{N}}$ were determined from the ^{15}N shifts in solution, since the pure 1-MeIm is a liquid and can not be measured in the solid-state.

From the crystalline $[\text{Mg-OEP}\cdot\text{H}_2\text{O}]$ and $[\text{Mg-OEP}\cdot\text{diazole}]$ complexes, IR spectra between 4000 and 300 cm^{-1} were recorded on a Perkin Elmer Paragon 1000 FTIR spectrometer. To observe the $\text{Mg-N}_4(\text{Por})$, $\text{Mg-O}_{\text{H}_2\text{O}}$ and $\text{Mg-N}_{\text{diazole}}$ vibration frequencies, far IR spectra of the crystalline $[\text{Mg-OEP}\cdot\text{H}_2\text{O}]$, the $[\text{Mg-OEP}\cdot 1,3\text{-}^{15}\text{N-Im}]$ and the $[\text{Mg-OEP}\cdot 1,3\text{-}^{15}\text{N}, 1,2\text{-}^{13}\text{C}\text{-}1\text{-MeIm}]$ complexes were recorded between 150 and 400 cm^{-1} on a Bruker 113v IR spectrometer.

5.3 Results

^1H NMR spectroscopy is used to determine the molar ratio of Mg-Por and ligand by the integrals of the ^1H responses and for an accurate determination of the ring current effect. Because it was expected that the carbon and nitrogen atoms are more changed by the electronic structure than the hydrogen atoms connected to the ligand, ^{15}N NMR and ^{13}C NMR experiments were applied to provide insight into the electronic structure of the axially coordinated ligand. In particular ^{15}N -NMR experiments were performed to deduce the electronic change of the nitrogen when it coordinates to the Mg-Por complex. IR spectra were recorded to complement and to confirm the NMR results.

5.3.1 ^1H -NMR

To determine the $\Delta\delta_{\text{H}}$, $\Delta\delta_{\text{C}}$ and $\Delta\delta_{\text{N}}$, first the isotropic chemical shifts of the free Im and 1-MeIm in CDCl_3 were determined. The $\delta_{\text{H}2}$, $\delta_{\text{H}4}$ and $\delta_{\text{H}5}$ of Im in CDCl_3 are 7.72, 7.02 and 7.02 ppm. For 1-MeIm, chemical shifts are measured for $\delta_{\text{H}2}$, $\delta_{\text{H}4}$, $\delta_{\text{H}5}$ and $\delta_{\text{CH}3}$ at 7.42, 6.87, 7.05 and 3.68 ppm. The $\delta_{\text{C}1}$ and $\delta_{\text{C}2}$ are observed at 137.9 and 33.3 ppm, respectively. The $\delta_{\text{C}2}$ of free $[1,3\text{-}^{15}\text{N}, 2\text{-}^{13}\text{C}]\text{-Im}$ in CDCl_3 is 134.9 ppm. The $\delta_{\text{N}1}$ and $\delta_{\text{N}3}$ of free $[1,3\text{-}^{15}\text{N}, 1,2\text{-}^{13}\text{C}]\text{-1-MeIm}$ are 161.2 and 256.1 ppm respectively. The individual $\delta_{\text{N}1}$ and $\delta_{\text{N}3}$ of $[1,3\text{-}^{15}\text{N}, 2\text{-}^{13}\text{C}]\text{-Im}$ can not be determined, because of the tautomerization that yields a $\bar{\delta}_{\text{N}1,\text{N}3} = 209.1$ ppm.

To resolve the $\Delta\delta_{\text{H}}$ of the ligand atoms, Mg-OEP and (B)Chl *a* are titrated with aliquots of Im or 1-MeIm solutions. As an example, NMR spectra of a titration series of 1-MeIm and BChl *a* are shown in figure 5.2. The maximum coordination shift ($\Delta\delta_{\text{H}}^{\text{max}}$), from the diamagnetic ring current effect of the porphyrin macrocycle is observed for $r_m \leq 1.0$. The $\Delta\delta_{\text{H}}$ decreases for $r_m > 1.0$ and vanishes for $r_m \rightarrow \infty$. The $\delta_{\text{H}5}$ and $\delta_{\text{CH}3}$ can be followed from the first added aliquot of ligand solution, since the resonances of the ligand are well resolved against the Mg-OEP or (B)Chl *a* background, while the responses from $\delta_{\text{H}2}$ and $\delta_{\text{H}4}$ are broadened and can not be resolved for $r_m < 1.0$. Both the line width and the $\Delta\delta_{\text{H}}$ are larger for H2 and H4 than for the other ligand hydrogens. For $r_m < 1.0$, the observed $\Delta\delta_{\text{H}5}$, $\Delta\delta_{\text{CH}3}$ are at their maximum and do not change when r_m is increased. For the $r_m > 1.0$, the $\Delta\delta_{\text{H}5}$ and

the $\Delta\delta_{\text{CH}_3}$ decrease after addition of each new aliquot of ligand. $\Delta\delta_{\text{H}2}$ and $\Delta\delta_{\text{H}4}$ appear as broad signals when $r_m > 4.3$. These signals narrow when the r_m is increased, while the ring-current shifts also vanish for $r_m \rightarrow \infty$.

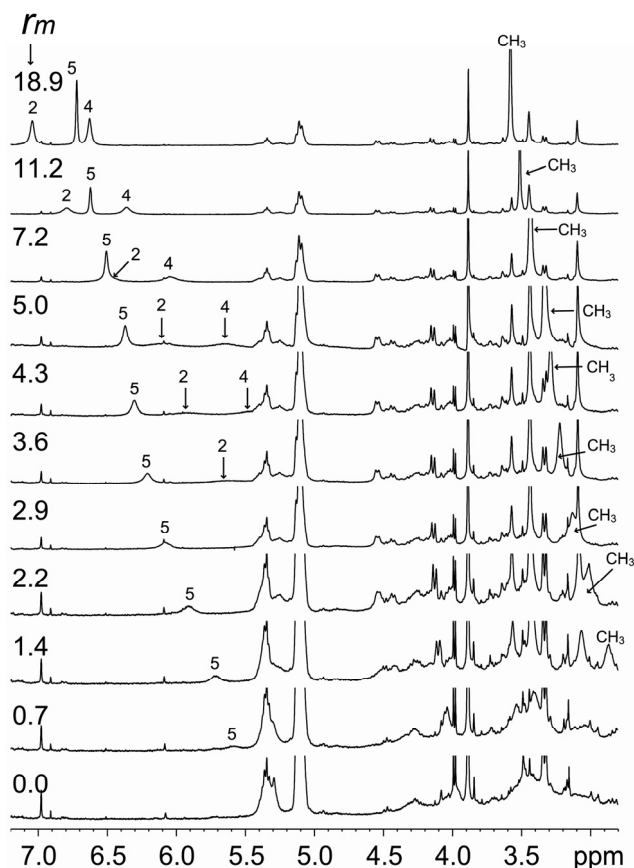


Figure 5.2 ^1H coordination shifts of 1-MeIm as function of $[\text{1-MeIm}]:[\text{BChl a}]$ molar ratios (r_m), measured in CDCl_3 at 25°C . The signals from $\delta\text{H}2$ and $\delta\text{H}4$ are indicated by arrows.

Evidently, an excess of Mg-Por leads to coordination of all available Im or 1-MeIm, leading to the maximum $\Delta\delta_{\text{H}}$. All titration experiments consistently show stronger ring-current effects for H2 and H4 than for H5 and CH_3 . While for the free Im the H4 and H5 resonances have the same isotropic shift due to the fast exchange between the two tautomeric forms, a different ring-current effect on H4

and H5 due to coordination makes them distinguishable. At a r_m of approximately 7.5 the δ_{H2} and the δ_{H5} are identical (figure 5.2). In all series, the crossing point of the δ_{H2} and the δ_{H5} is found at ~ 6.5 ppm. The $\Delta\delta_H^{\max}$ for Im and 1-MeIm in the titration series with Mg-OEP or (B)Chl *a* as a ring-current reagent are summarized in table 1. For the [(B)Chl *a*-diazole] complexes the $\Delta\delta_{H2}^{\max}$ and $\Delta\delta_{H4}^{\max}$ are not observed due to broadening of the signals or overlap with the (B)Chl *a* porphyrin background signal. In contrast, for the relatively simple [Mg-OEP-diazole] complexes, a $\Delta\delta_H^{\max}$ is observed for all ligand hydrogens.

For the titration series with Mg-OEP the $\Delta\delta_{H2}^{\max}$, $\Delta\delta_{H4}^{\max}$, $\Delta\delta_{H5}^{\max}$ and $\Delta\delta_{CH3}^{\max}$ are all observed and the titration curves of the $\Delta\delta$ can be scaled by setting $\Delta\delta_H^{\max} = 1.0$. As an example, figure 5.3A shows the scaled titration curves of $\Delta\delta_{H2}$, $\Delta\delta_{H4}$, $\Delta\delta_{H5}$ and $\Delta\delta_{CH3}$ of the 1-MeIm for various r_m . The four $\Delta\delta$ curves are comparable. This reflects that the variation of $\Delta\delta$ with r_m is due to the same molecular events for all hydrogens. Similar behavior is observed for the $\Delta\delta_{H2}$, $\Delta\delta_{H4}$ and $\Delta\delta_{H5}$ of the Im ligand coordinated to Mg-OEP. For the (B)Chl *a* series, only $\Delta\delta_{H5}$ and $\Delta\delta_{CH3}$ curves can be compared, since the $\Delta\delta_{H2}^{\max}$ and $\Delta\delta_{H4}^{\max}$ are not observed for $r_m < \sim 4.0$. However, when the observable parts of the $\Delta\delta_{H2}$ and $\Delta\delta_{H4}$ curves, for $r_m > \sim 4.0$ are scaled also, a good correlation with the $\Delta\delta_{H5}$ and $\Delta\delta_{CH3}$ normalized data is obtained. Subsequently, the $\Delta\delta_{H2}^{\max}$, $\Delta\delta_{H4}^{\max}$ were estimated by extrapolation. For the titration series in which Im was added to (B)Chl *a*, no full $\Delta\delta_{H5}^{\max}$ titration curve was observed, since the $\Delta\delta_{H5}$ in [(B)Chl *a*-Im] complexes is broadened for $r_m < 1.0$. In this case, it is not possible to scale the $\Delta\delta_{H5}^{\max}$ to 1.0, which is necessary for the extrapolation of the $\Delta\delta_{H2}$ and $\Delta\delta_{H4}$ curves.

The $\Delta\delta_{H_{meso}}$ (H5, H10, H15 and H20 of the Mg-OEP) shows a small significant gradual increase of the $\Delta\delta_{H_{meso}}$ for $r_m < 1.0$ and reaches the $\Delta\delta_{H_{meso}}^{\max}$ at $r_m = 1.0$. For $0.0 < r_m \leq 1.0$, the $\Delta\delta_H$ of the ligand is maximally and does not change. The increase of the $r_m > 1.0$, yields a decrease of the $\Delta\delta_H$ of the ligand hydrogens (figure 5.3A), while the $\Delta\delta_{H_{meso}}^{\max}$ is essentially constant for $r_m > 1.0$ (figure 5.3B). The ethyl groups of Mg-OEP are hardly affected by the axial coordination. A $\Delta\delta_{H_{meso}}^{\max}$ for $\Delta\delta_{H5}$, $\Delta\delta_{H10}$ and $\Delta\delta_{H20}$ of the pure [(B)Chl *a*-diazole] complex dissolved in CDCl₃ can not be derived, since the δ_{H5} , δ_{H10} and δ_{H20} signals are broadened beyond the detection limit due to aggregation of the chlorophyll molecules. When a small amount of ligand is

added, broad δ_{H5} , δ_{H10} and δ_{H20} signals are observed that are not shifted. The δ_{H5} , δ_{H10} and δ_{H20} signals narrow upon increasing r_m to $r_m = 1.0$.

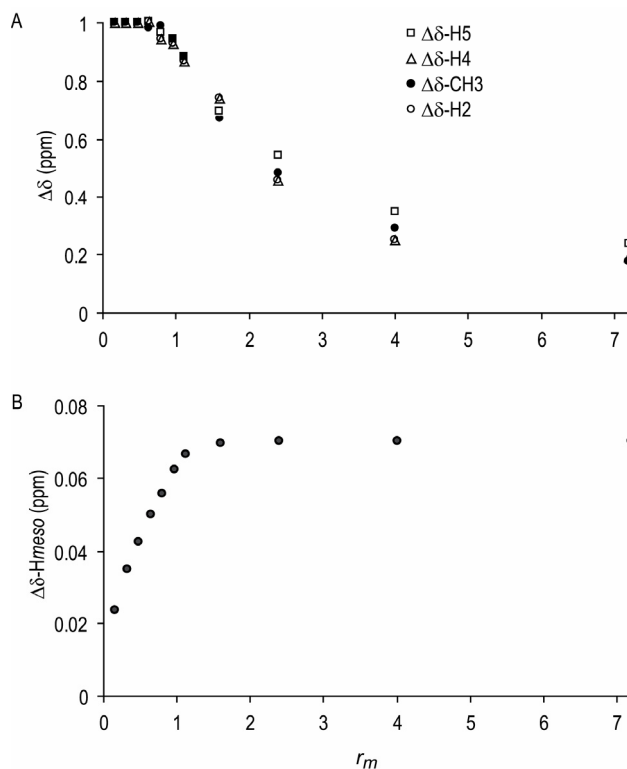


Figure 5.3 A) Normalized $\Delta\delta H$ of the ligand hydrogens in the $[Mg(II)\text{-OEP}\cdot 1\text{-MeIm}]$ complex for various r_m , measured in $CDCl_3$ at 25 °C. B) $\Delta\delta H_{meso}$ of the $[Mg(II)\text{-OEP}\cdot 1\text{-MeIm}]$ for various r_m 's, measured in $CDCl_3$ at 25 °C.

5.3.2 ^{13}C -NMR and ^{15}N -NMR

The $\Delta\delta_{C2}^{\max}$ and $\Delta\delta_{CH3}^{\max}$ for the $[1,3\text{-}^{15}N, 2\text{-}^{13}C]\text{-Im}$ and $[1,3\text{-}^{15}N, 1,2\text{-}^{13}C]\text{-1-MeIm}$ in complex are summarized in table 1. The coordination effects on the $\Delta\delta_C$, due to ring currents and electronic perturbations when Im or 1-MeIm coordinates to the Mg, are consistent with the effects measured for the $\Delta\delta_H$. Also a gradual decrease of the $\Delta\delta_C$ is observed for $r_m > 1.0$.

Table 5.1 The coordination shift $\Delta\delta H^{\max}$, $\Delta\delta C^{\max}$ and $\Delta\delta N^{\max}$ (ppm) of imidazole and 1-methyl imidazole coordinated to Mg-OEP and (B)Chl *a*.

Im-complex	H ₂	H ₄	H ₅	NH	H-meso	C ₂
Mg-OEP	4.9	4.7	2.3	4.2	0.08	4.5
Chl <i>a</i>			1.6 ^a			
BChl <i>a</i>			1.0 ^a			1.8

a) Observable for a $r_m > \pm 2.5$

1-Melm-complex	H ₂	H ₄	H ₅	CH ₃	H-meso	C ₂	CH ₃	N ₁	N ₃
Mg-OEP	5.2	5.2	2.3	1.6	0.07	5.5	1.3	2.1	27.2
Chl <i>a</i>	3.0 ^b	3.0 ^b	1.7	1.1					
BChl <i>a</i>	2.8 ^b	2.7 ^b	1.5	0.9		2.2	0.2		

b) Extrapolated values.

For the ¹⁵N NMR measurements, the HMBC technique that provides the ¹H-¹⁵N correlations, was performed on the crystalline [Mg-OEP-1,3-¹⁵N, 1,2-¹³C-1-MeIm] and Mg-OEP·1,3-¹⁵N-Im] complexes dissolved in anhydrous CDCl₃. The r_m is exactly 1.0, pointing to a single axial coordination of the Mg-Por complex. The $\Delta\delta_H^{\max}$ of the ligand hydrogens can be detected in the ¹H NMR spectrum and are in line with the $\Delta\delta_H^{\max}$ obtained from the titration series. The HMBC NMR spectrum of the [Mg-OEP·1,3-¹⁵N, 1,2-¹³C-1-MeIm] complex is presented in figure 5.4. Strong signals are observed for the ¹⁵N1/CH₃ and the ¹⁵N1/H5 correlation, while weaker correlation signals are observed for the ¹⁵N1/H2, ¹⁵N3/H2, ¹⁵N3/H4, and the ¹⁵N3/H5 cross resonances. The 2D NMR spectrum also shows a correlation between the $\delta_{H_{meso}}$ at 10.1 ppm and the natural abundance signal from the four N nuclei of Mg-OEP at $\delta_N = 200$ ppm. The ¹⁵N₁ and ¹⁵N₃ are detected with $\delta_{N1} = 159.1$ and $\delta_{N3} = 230.9$ ppm, corresponding with a $\Delta\delta_N^{\max}$ of 2.1 and 25.2 ppm, respectively. This is also shown in figure 5.5D. The strong ¹J_{H¹³C couplings between H2 and ¹³C2 (206}

Hz) and between $^{13}\underline{\text{C}}\text{H}_3$ and $^{13}\underline{\text{C}}\text{H}_3$ (140 Hz) confirm the assignment of both nitrogen signals. When the r_m is increased to 2.5, the $\delta_{\text{N}3}$ broadens beyond the detection limit. This indicates exchange of the coordinated 1-Melm with the 1-Melm pool in solution. For $r_m > 10$ the ^{15}N NMR spectrum is essentially identical to the spectrum observed from the free ligand solution.

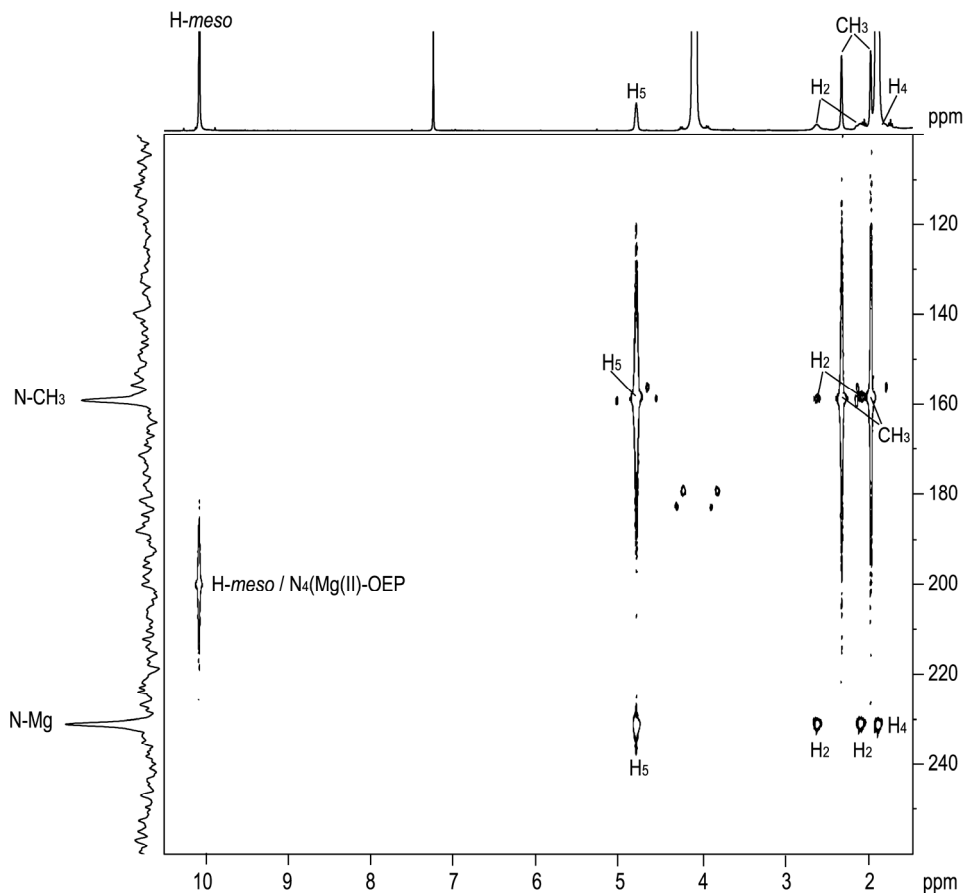


Figure 5.4 ^1H - ^{15}N HMBC NMR correlation data set of micro-crystalline $[\text{Mg}(\text{II})\text{-OEP}\cdot[1,3\text{-}^{15}\text{N}, 1,2\text{-}^{13}\text{C}]\text{-1-Melm}]$ dissolved in CDCl_3 . The large $\Delta\delta\text{H}2^{\text{max}}$, $\Delta\delta\text{H}4^{\text{max}}$ and $\Delta\delta\text{H}5^{\text{max}}$ and the relative signal intensity in the ^1H NMR in combination with the correlation between H2, H4, H5 and the $^{15}\text{N}1$ at 231 ppm, prove the existence of the five-coordinated $[\text{Mg}(\text{II})\text{-OEP}\delta\text{-}1\text{-Melm}\delta]^+$ complex.

The HMBC NMR spectrum of the [Mg-OEP-1,3-¹⁵N-Im] complex (data not shown) clearly shows the ¹⁵N1/NH1 ($^1J_{\text{H}^{15}\text{N}} = 96$ Hz) and the ¹⁵N1/H5 correlation. The correlation between the $\delta_{\text{H}_{\text{meso}}}$ at 10.1 ppm and the natural abundance δ_{N} of the Mg-OEP at 200 ppm is also observed. The correlations ¹⁵N3/H2, ¹⁵N3/H4, and ¹⁵N3/H5 are too weak to be detected. The 1D solution ¹⁵N NMR spectrum of the [Mg-OEP-1,3-¹⁵N-Im] complex in figure 5.5B shows two ¹⁵N signals at 232.9 and 154.0 ppm from the N3-Mg and N1H ($^1J_{\text{H}^{15}\text{N}} = 96$ Hz) respectively. These isotropic chemical shifts are comparable to the ¹⁵N shifts measured for the [Mg-OEP-1,3-¹⁵N, 1,2-¹³C-1-MeIm] complex. The averaged δ_{N1} and δ_{N3} of free [1,3-¹⁵N]-Im in solution makes it impossible to determine the $\Delta\delta_{\text{N}}$ in solution.

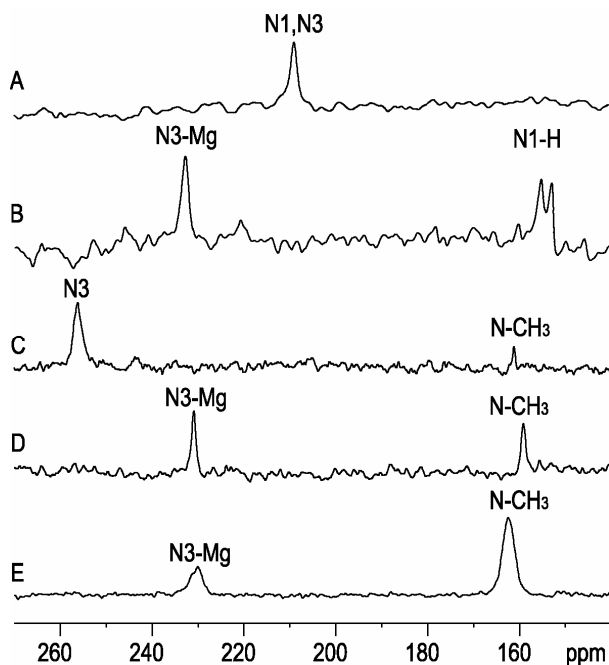


Figure 5.5 ¹⁵N NMR spectra of the free [1,3 ¹⁵N]-Im (A), the monomeric [Mg-OEP-1,3 ¹⁵N-Im] complex, $^1J_{\text{H}^{15}\text{N}} = 96$ Hz (B), the free [1,3 ¹⁵N, 1,2-¹³C]-1-MeIm (C), the monomeric [Mg-OEP-1,3 ¹⁵N, 1,2 ¹³C-1-MeIm] complex in CDCl₃ (D), and the MAS NMR spectrum of the crystalline [Mg(II)-OEP-1,3 ¹⁵N, 1,2 ¹³C-1-MeIm] complex (E).

Figure 5.5E shows the 1D MAS NMR spectrum of the [Mg-OEP·1,3-¹⁵N, 1,2-¹³C-1-MeIm] complex. The measurement demonstrates the existence of the model complex in the solid-state, and that it has exactly one axial ligand, forming a five-coordinated complex, which corresponds with the (B)Chl α -his(Im) complex natural photosynthetic complexes.¹⁶ The $\Delta\delta_N$ provided by the MAS ¹⁵N NMR spectra of the [Mg-OEP·1,3-¹⁵N, 1,2-¹³C-1-MeIm] complex are in close agreement with the shifts observed for the monomeric complex in solution. The δ_{N3} (N3-Mg) appears at 229 ppm and corresponds with $\Delta\delta_N = 27.2$ ppm. The $\Delta\delta_{N1}$ is only 0.8 ppm.

5.3.3 IR spectra

To complement the NMR data, a few IR data sets from the complexes have been recorded. The OH stretching vibration in the IR spectrum of [Mg-OEP·H₂O] at 3529 cm⁻¹ (data not shown) has disappeared in the spectra of the [Mg-OEP·1,3-¹⁵N-Im] and [Mg-OEP·1,3-¹⁵N, 1,2-¹³C-1-MeIm] complexes. This shows that a water molecule is substituted by Im and 1-MeIm, respectively. In the IR spectrum of the [Mg-OEP·1,3-¹⁵N-Im] complex a response is observed at 3347 cm⁻¹, which is assigned to the N-H stretching vibration of the axially coordinated Im.

The low-frequency regions in the far infra red spectra of all three complexes are presented in figure 5.6. It shows a vibration at 350 cm⁻¹ with a shoulder at ~340 cm⁻¹. These vibrations are assigned to the symmetric and asymmetric Mg-N₄(Por) stretching vibration, respectively. Upon substitution of the axial water ligand by Im or 1-MeIm, the Mg-O_{H2O} stretching vibration at 215 cm⁻¹ disappears and a new response from the Mg-N_{1-MeIm} or Mg-N_{Im} stretching vibration can be observed at 300 cm⁻¹, pointing to a dynamic complex.

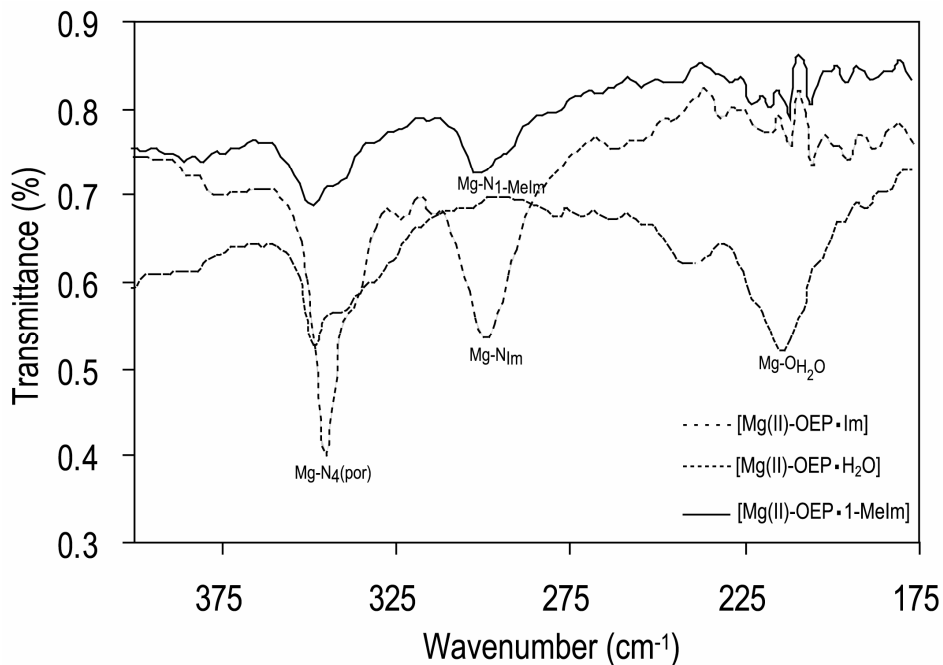


Figure 5.6 Low-frequency IR spectrum of the $[Mg(II)-OEP\cdot H_2O]$ complex (dashed line), the crystalline $[Mg(II)-OEP-1,3^{15}N, 1,2^{13}C-1-Melm]$ complex (solid line) and the $[Mg(II)-OEP-1,3^{15}N-Im]$ complex (grey line).

5.4 Discussion

Monomeric (B)Chl, dissolved in polar solvents, shows very sharp and narrow resonances in the 1H NMR spectra.¹⁷ (B)Chl *a* in $CDCl_3$ is aggregated, and does not form solvate complexes.^{18,19} In our model study this is reflected by the broadened signals of the H_5 , H_{10} , and H_{20} *meso*-hydrogens ($\delta_{H_{meso}}$) from pure (B)Chl *a* in $CDCl_3$. The δ_{H_5} , $\delta_{H_{10}}$ and $\delta_{H_{20}}$ responses of (B)Chl *a* show up gradually and narrow when Im or 1-Melm is added to the (B)Chl *a* solution. This shows that axial ligand coordination is energetically favorable compared to maintaining an aggregate structure. For $r_m \geq 1.0$, the H_{meso} resonances are narrow, and do not show an upfield shift due to neighboring porphyrins like this occurs in aggregate structures. This provides evidence for the formation of the [(B)Chl *a*:diazole] complex in $CDCl_3$.

The $\delta_{H_{meso}}$ of the [Mg-OEP·H₂O] and the $\delta_{H_{meso}}$ of the [Mg-OEP·diazole] complexes are narrow for both $r_m < 1.0$ and $r_m \geq 1.0$ and are not shifted due to the aggregation effects, which gives convincing evidence for monomeric [Mg-OEP·H₂O] and [Mg-OEP·diazole] complexes in CDCl₃.

A constant $\Delta\delta_H$ in all titration experiments at $r_m < 1.0$ shows that virtually all added ligand molecules are coordinated to the Mg ion. The decrease of the $\Delta\delta_H$ for $r_m > 1.0$ is attributed to a rapid exchange of coordinated ligand with molecules from the ligand pool, since in that case averaging of the isotropic shifts of the coordinated and free ligand occurs, according to their relative concentrations (figure 5.2).

The $\delta_{H_{meso}}$ for $r_m < 1.0$ is due to rapid averaging of the responses from the coordinated and bare porphyrin. In that case, all the available ligand is coordinated and the excess of Mg-Por exchanges with the coordinated form. At $r_m = 1.0$ the $\Delta\delta_{H_{meso}}^{max}$ is reached, while at this point the decrease of the $\Delta\delta_H$ of the ligand starts. This means a minimum of exchange at $r_m = 1.0$.

Since the $\Delta\delta_{H_{meso}}^{max}$ is reached for $r_m \geq 1.0$ (figure 5.3B), it can be concluded that the Mg-Por in solution complexes are five-coordinated. The binding of N-donor ligands in solution to Mg-Por, like pyridines and diazoles, has been studied thoroughly and it is clear that the dominant process in solution and in most crystal structures is the formation of five-coordinate 1:1 adducts that have a puckered structure.^{20,21} Axial coordination to the Mg of (B)Chl *a* has been investigated for monomeric BChl *a* complexes in solution by resonance Raman spectroscopy, magnetic circular dichroism, and absorption spectroscopy.^{18,19,22,23} These studies confirm that the Mg in (B)Chl is mostly five-coordinated in aggregates and in solution with imidazole derivatives.

The $\Delta\delta_H^{max}$ of the [Mg-OEP·1-MeIm] and [Mg-OEP·Im], summarized in table 1, show that $\Delta\delta_H^{max}$ of all ligand hydrogens is slightly larger for 1-MeIm than for Im, which is in line with a stronger coordination effect for the 1-MeIm than for Im. In line with the $\Delta\delta_H$ results, the larger $\Delta\delta_{C2}$ for the [Mg-OEP·1-MeIm] complex than for the [MgPor·Im] complex confirms the stronger coordination effect for 1-MeIm than for Im. This stronger coordination is a result of the methyl group, which is a better electron-donating group than a hydrogen atom.

The C2 in the diazole ring is probably more sensitive to variations of the electronic structure due to coordination than the hydrogens that are connected to the ring. However, the $\Delta\delta_{\text{C2}}$ and $\Delta\delta_{\text{CH3}}$ both decrease when the $r_m > 1.0$, very similar to the $\Delta\delta_{\text{H}}$ in the titration series. This shows that the $\Delta\delta_{\text{C}}$ is to a large extent determined by the ring current. The $\Delta\delta_{\text{C2}}^{\text{max}}$ in our experiments is 4.5 and 5.3 ppm for coordinated Im and 1-MeIm to Mg-OEP respectively, while the $\Delta\delta_{\text{C2}}$ of Im and 1-MeIm coordinated to BChl *a* are 1.8 and 2.2 ppm, respectively. The data for the C2 in the [Mg-OEP-diazole] models match very well with $\Delta\delta_{\text{C}} = 5.1$ ppm for the corresponding C ϵ of the histidyl residue in the natural LH2 complex.¹⁶ This shows that the Mg-N3 coordination in the natural system is stronger than in the [BChl *a*-1-MeIm] model complex. In this regard, the stabilizing effect on the coordination bond of the methyl group at the N π in our model system provokes a comparable effect for the protein environment on the stabilization of the Mg-N coordination bond in the natural system. The protein potentially has an electron donating effect on the NH1 moiety.

The intermolecular nuclear shielding is a probe for the distance from the nucleus to the plane of the macrocycle.²⁴ The ring current effects depend more strongly on the distance to the plane of the macrocycle for nuclei at a short distance than for positions that are more remote from the equatorial plane.²⁴⁻²⁶ Therefore, the $\Delta\delta$ for nuclei closer to the porphyrin ring are more sensitive to variations of the distance than the $\Delta\delta$ of nuclei that are remote. Due to vibrations of the Mg-N bond as found by the IR measurements, the distance between the nuclei and the macrocycle slightly changes continuously, yielding a broadened response of the ligand nuclei. This is reflected in the line widths, which are larger for the proximal H2 and H4 than for the distal H5 and the CH₃. From the equal $\Delta\delta_{\text{H2}}^{\text{max}}$ and $\Delta\delta_{\text{H4}}^{\text{max}}$ (table 1), it is concluded that the distances between H2 and the macrocycle and between H4 and the macrocycle are equal. As a consequence of the symmetry of the ligand, the distances between H5 and the macrocycle and between NH1 and the macrocycle are also equal. It is remarkable that the H1 in the [Mg-OEP-Im] with $\Delta\delta_{\text{NH1}}^{\text{max}} = 4.2$ ppm is more shielded than the H5 with $\Delta\delta_{\text{H5}}^{\text{max}} = 2.3$ ppm, while the distances for H5 and NH1 to the Mg should be comparable, according to electronic calculations performed on this model system, which show that an identical shielding effect can be expected for both H5 and NH1 (unpublished results). Moreover, the $\Delta\delta_{\text{NH1}}$

titration curve can not be scaled to the normalized curves of $\Delta\delta_{\text{H}2}$, $\Delta\delta_{\text{H}4}$, and $\Delta\delta_{\text{H}5}$. This indicates that the $\Delta\delta_{\text{NH}1}$ can not be only a result of the ring current effect and involves another mechanism as well. The $\text{NH}1$ signal of free Im in solution is observed at 10.8 ppm, which is significantly higher than the calculated value of 8.1 ppm, while all other calculated shifts correspond with the observed values in the blank experiment. The relatively high $\delta_{\text{NH}1}$ of Im in solution is attributed to intermolecular hydrogen-nitrogen interactions of Im molecules, forming loose aggregates, while the calculations are performed on a monomer. For the [Mg-OEP·Im] complex, the calculated $\delta_{\text{NH}1} = 6.1$ and the observed $\delta_{\text{NH}1} = 6.6$ are comparable, which confirms that $\delta_{\text{NH}1}$ of Im in solution is anomalous.

The HMBC NMR spectrum of the dissolved crystalline [Mg-OEP·1,3- ^{15}N , 1,2- ^{13}C -1-MeIm] complex, presented in figure 5.4, shows that the exchange processes are quenched exactly at the equimolar ratio when $r_m = 1.0$ and a stable complex is obtained in solution. This is in line with the ^1H -NMR spectra. Both the pyrrole and the pyridine type ^{15}N are observed, while the pyridine type ^{15}N can not be observed in the ^{15}N NMR data sets of the titration series due to exchange broadening. The $\delta_{\text{N}3}$ (N-Mg) at 230.9 ppm in the HMBC NMR spectrum of the [Mg-OEP·1,3- ^{15}N , 1,2- ^{13}C -1-MeIm] complex in CDCl_3 and at 228.9 ppm in the ^{15}N MAS NMR data set implies a $\Delta\delta_{\text{N}}$ of 25.2 - 27.2 ppm. Such a large $\Delta\delta_{\text{N}}$ can not be produced by ring currents, and provides convincing evidence for a strongly perturbed electronic structure. The shift of the $\delta_{\text{N}3}$ (N-Mg) is close to the $\delta = 224.0$ ppm observed for the $\delta_{\text{N}3}$ (N-Mg) in the LH2 antenna complex.⁶ The large $\Delta\delta_{\text{N}}^{\text{max}}$ of the N3 (N-Mg) possibly points to a partial charge transfer from the 1-MeIm to the Mg-OEP, yielding a $\text{Mg-OEP}^{\delta-}/1\text{-MeIm}^{\delta+}$ complex, which was observed for the natural LH2 complex.¹⁶ In a more general perspective, the coordination of histidine and imidazole in the model compounds are also of interest for the studies of other porphyrin or heme containing proteins that have dynamic axial coordination properties to perform their biological key function.

The $\Delta\delta$ is a useful marker to provide insight into the axial coordination of diazole ligands coordinated to Mg-porphyrins. The $\Delta\delta_{\text{H}}$ data compare well with previous results from other Mg-porphyrin systems that also show a gradual change of the $\Delta\delta_{\text{H}}$ after increasing the methanol or pyridine concentrations.^{17,27} In all cases, the NMR measurements were done in a non-polar solvent.

Finally, in the NMR study on the [(B)Chl *a*-Im] complex in THF relatively small shifts (< 0.5 ppm) for the coordinated Im hydrogens and no coordination was observed when 1-MeIm was codissolved with (B)Chl *a* in THF.⁸ In this study, a negatively charged [(B)Chl *a*-Im] complex was inferred to explain the results. In the present study, a single neutral axial coordination for both the [Mg-Por-1-MeIm] and [Mg-Por-Im] complex in CDCl₃ with large hydrogen shifts for both coordinated Im and 1-MeIm has been observed.

5.5 Conclusion

The use of advanced chemical models to understand the molecular structure at the molecular level has been clearly demonstrated. The synthetic analogues of the histidine-(bacterio)chlorophyll complex can mimic their structural features and electronic properties for a better understanding of the natural system, and the information from these models is surely also of interest for other metal-containing complexes like the heme proteins and vitamin B12, in which the dynamic coordination bond is a prerequisite for their functionality.

References

- (1) Scheer, H.; *The Chlorophylls*. CRC Press: Boca Raton FL, 1991.
- (2) Branden, C.; Tooze, J.; In: *Introduction to protein structure.*; Garland Publishing, Inc.: New York, 1991.
- (3) Robert, B.; Lutz, M.; *Biochim et Biophys. Acta* 1987, 807, 10-23.
- (4) Prince, S. M.; Papiz, M. Z.; Freer, A. A.; McDermott, G.; Hawthornthwaite-Lawless, A. M.; Cogdell, R. J.; Isaacs, N. W.; *J. Mol. Biol.* 1997, 268, 412-423.
- (5) Storm, C. B.; *J. Am. Chem. Soc.* 1970, 92, 1423-1425.
- (6) Soede-Huijbrechts, C.; Cappon, J. J.; Boender, G. J.; Gast, P.; Hoff, A. J.; Lugtenburg, J.; de Groot, H. J. M.; In: *Photosynthesis: mechanisms and effects.*; Kluwer Academic Publ.: Dordrecht, 1998.
- (7) Sham, Y. Y.; Muegge, I.; Warshel, N.; *Biophys. J.* 1999, 76, A198-A198.

- (8) Alia; Matysik, J.; Erkelens, C.; Hulsbergen, F. B.; Gast, P.; Lugtenburg, J.; de Groot, H. J. M.; *Chem. Phys. Lett.* 2000, 330, 325-330.
- (9) Gridnev, A. A.; Mihaltseva, I. M.; *Synth. Commun.* 1994, 24, 1547-1555.
- (10) Withanowski, M.; Webb, G. A.; In: *Nitrogen NMR 1973*, Plenum Press London and New York. 205.
- (11) Chen, B. B.; von Philipsborn, W.; Nagarajan, K.; *Helv Chim. Acta* 1983, 66, 1537-1555.
- (12) Welleman, J. A.; Hulsbergen, F. B.; Verbiest, J.; Reedijk, J.; *J. Inorg. Nucl. Chem.* 1978, 40, 143-147.
- (13) Lindsey, J. S.; Woodford, J. N.; *Inorg. Chem.* 1995, 34, 1063-1069.
- (14) Omata, T.; Murata, N.; *Photochem. Photobiol.* 1979, 31, 183-185.
- (15) EgorovaZachernyuk, T. A.; van Rossum, B.; Boender, G. J.; Franken, E.; Ashurst, J.; Raap, J.; Gast, P.; Hoff, A. J.; Oschkinat, H.; de Groot, H. J. M.; *Biochemistry* 1997, 36, 7513-7519.
- (16) Alia; Matysik, J.; Soede-Huijbregts, C.; Baldus, M.; Raap, J.; Lugtenburg, J.; Gast, P.; van Gorkom, H. J.; Hoff, A. J.; de Groot, H. J. M.; *J. Am. Chem. Soc.* 2001, 123, 4803-4809.
- (17) Katz, J. J.; Strain, H. H.; Leussing, D. L.; Dougherty, R. C.; *J. Am. Chem. Soc.* 1968, 90, 784-791.
- (18) Nozawa, T.; Ohtomo, K.; Suzuki, M.; Morishita, Y.; Madigan, M. T.; *Bull. Chem. Soc. Jpn.* 1993, 66, 231-237.
- (19) Umetsu, M.; Wang, Z. Y.; Kobayashi, M.; Nozawa, T.; *Biochim. Biophys. Acta-Bioenerg.* 1999, 1410, 19-31.
- (20) Dimagno, S. G.; Lin, V. S. Y.; Therien, M. J.; *J. Am. Chem. Soc.* 1993, 115, 2513-2515.
- (21) Taylor, P. N.; Wylie, A. P.; Huuskonen, J.; Anderson, H. L.; *Angew. Chem.-Int. Edit.* 1998, 37, 986-989.
- (22) Umetsu, M.; Wang, Z. Y.; Yoza, K.; Kobayashi, M.; Nozawa, T.; *Biochim. Biophys. Acta-Bioenerg.* 2000, 1457, 106-117.
- (23) Fujiwara, M.; Tasumi, M.; *J. of Phys. Chem.*, 1986; Vol. 90, 250-255.
- (24) Giessner-Prettre, C.; Pullman, B.; *J. Theor. Biol.* 1971, 31, 287-294.
- (25) Abraham, R. J.; Bedford, G. R.; McNeillie, D.; Wright, B.; *Org. Mag. Res.* 1980, 14, 418-425.
- (26) Gouedard, M.; Gaudemer, F.; Gaudermer, A.; Riche, C.; *J. Chem. Res. (S)* 1978, 30-31.
- (27) Brereton, R. G.; Sanders, J. K. M.; *J. Chem. Soc. Trans. I*; 1983, 423-237.

Chapter 6

General discussion and prospects

6.1 Sequence specific assignments of α -helical transmembrane proteins

The main target of the research described in this thesis is the sequence specific backbone and side chain assignment for the α -helical constructed LH2 protein and a selective assignment for the three bacteriochlorophyll cofactors embedded in the monomeric LH2 unit of the nine-membered aggregated complex. The chemical shift assignment of the protein with MAS NMR spectroscopy is a first essential step of the structure determination procedure.¹ Resolving chemical shifts of α -helical constructed proteins, such as the LH2 protein, is difficult because of the narrow shift distribution of the backbone nuclei in the spectra. This thesis describes the development of biosynthetic pattern labeling as a method for resolving chemical shifts of backbone nuclei of α -helical protein constructs. The biosynthetic labeling method was developed to reduce the number of correlations in the spectra and has appeared successful for resolving chemical shift data of backbone nuclei in the LH2

transmembrane protein. The sequence specific assignment for 76 residues of the 94 residue containing monomeric LH2 protein unit was realized by applying conventional more-dimensional MAS NMR techniques on novel pattern labeled samples, which has been described in chapter 2 and chapter 3. It has been shown that biosynthetic pattern labeling in combination with conventional high-resolution solid-state NMR methods can help to resolve chemical shifts and leads to sequence specific assignments of α -helical transmembrane protein structures. The ultimate goal of the study on the LH2 complex will be to determine the protein structure of the predominantly α -helical constructed protein complex by solid-state NMR, including the determination of the quaternary structure of the complex. In the NMR analysis presented in chapter 3, some contacts between remote residues and cofactor-protein correlations are observed. Inter-residue interactions are observed between residues within a single turn of a α -helix or for residues within a loop segment. This provides some secondary structure information. To resolve tertiary structure it is essential to detect inter-helix contacts. Since the distance between the helical rods in the LH2 protein complex is about 9.0 Å, it is difficult to obtain inter-helix correlations between backbone nuclei. This contrasts with the β -sheet constructed proteins where inter-chain correlations between backbone carbons are readily detected due to hydrogen bonding with the adjacent chain in the β -pleated sheet.

Embedded between the transmembrane helices of the α - and β -subunits of a monomeric LH2 unit are the BChl *a* cofactors. Since the helices on both sides of the cofactors belong to different subunits, the protein-cofactor interactions, described in chapter 3, provide information about the structural organization of the subunits, which is the quaternary protein structure, while the tertiary structure is still unknown. The next step towards the structure determination of the LH2 3D structure by MAS NMR will be to collect inter-residue correlations between residues from adjacent transmembrane helices. In the helical segments, the carbonyl groups in the protein backbone are hydrogen bonded with residues at $i\pm 3$ or $i\pm 4$ and the side-chains of the residues are directed away from the helical rod. The closest contacts between the adjacent helices are between the ends of the side-chains of the residues. To resolve inter-helix contacts between aliphatic side-chains, a focus on correlations between methyl groups of aliphatic residues is recommended. These

are the closest contacts between helices and abundant in α -helical transmembrane proteins.

In solution NMR much of the success in structure determination relates to the high abundance of short ^1H - ^1H distances that characterize the 3D structure of the protein. While in solid-state NMR strong dipolar couplings between the ^1H spins are problematic for a high-resolution, these couplings can be used for an effective transfer of the polarization. Recently, the CP³ 2D CHHC MAS NMR correlation experiment with mixing by true ^1H spin diffusion was presented.^{2,3} This method exploits the strong dipolar couplings in the network of ^1H spins to enable the observation of intermolecular ^1H - ^1H contacts in aggregated chlorophylls detected by the ^{13}C response. It has been demonstrated that the CHHC experiment can also be used to provide structural restraints for the folding of a protein.^{4,5} For protein samples, the ^{13}C - ^{13}C spin diffusion spectra are dominated by intraresidue correlations, since the coupling network of the ^{13}C spins is mostly restricted to direct bound ^{13}C neighbors. This contrasts with the topology of ^1H spins, which form a widely branched network throughout the protein with both strong intraresidual and strong interresidual couplings. Hence the CHHC experiment is suitable for resolving inter-helical correlations via the many short distance interhelical ^1H - ^1H contacts between aliphatic side chains from different transmembrane protein domains in the protein. The pattern labeled samples can be very helpful to collect CHHC distance restraints. From figure 2.4 can be deduced that the 1.4-LH2 sample can be used for resolving contacts between Thr, Asp, Asn, Glu, Gln, Arg and His residues, while the 2.3-LH2 sample is important for the detection of contacts between Ser, Pro, Ala, Val and K residues. The 2.3-LH2 sample is potentially also useful for the observation of aliphatic contacts between the cofactor and the protein, since all side groups of the BChl *a* in this sample are labeled.

6.2 Computational methods

In the light of several ongoing structural genomics projects, fast and reliable methods for structure determinations from NMR data are in great demand. It has

been demonstrated in chapter 2 and 3 that the pattern labeling approach is very useful to accelerate the assignment procedure. Simultaneously, a computational aid at an early stage of the structure determination procedure is more than desirable. “Torsion Angle Likelihood Obtained from shifts and sequences Similarity (TALOS) is a frequently used computer program, which has been developed to search the Brookhaven protein database and the BioMagRes Bank for strings of residues with residue type and chemical shift homology, respectively.⁶ TALOS reads the experimental protein chemical shifts and converts them into secondary chemical shifts before entering them in the database. Chemical shifts of backbone nuclei in proteins are sensitive to local conformation, and homologous proteins show similar patterns of secondary chemical shifts. The inverse of this relation is used to search a database for triplets of adjacent residues with secondary chemical shifts and sequence similarity that provide the best match to the query triplet of interest. TALOS searches for the 10 triplets with the closest match of secondary shifts and amino acid sequence to the query sequence. Subsequently, the probability maps can be incorporated into the structure calculation methods for both large and small proteins.

In a first attempt, TALOS was applied to the shift data in tables 3.1a and 3.1b of chapter 3 for a prediction of the structural elements of the triplets. The correspondence of the predicted structures with the structural data from the high-resolution X-ray structure was poor. The dissimilarity may be attributed to the presence of the BChl *a* cofactors, which may affect the details of the electronic structure of the protein. In addition the macrocycles of the B800 and B850 cofactors induce ring current shifts at nearby nuclei, which are not yet handled by the TALOS program.

For solution NMR, programs for automated Overhauser effect (NOE) peak assignments have been developed.⁷⁻¹³ One of the successful programs is Ambiguous Restraints for Iterative Assignment (ARIA).¹³ It accelerates the NOE assignment by the reduction of ambiguous distance restraints in an iterative structure calculation strategy. By comparing theoretical and measured NOE intensities, ARIA estimates target distances for the next iteration of structure calculation. To some extent, the correlations observed in CHHC MAS NMR spectra can be compared with nuclear NOE peaks in solution NMR. A manual NOE assignment of α -helices is difficult due

to the strong overlap in the spectra. It is expected that this will be similar for MAS CHHC NMR experiments applied on uniformly labeled membrane protein samples in the solid state. The samples with reduced labeling may contribute here to resolve CHHC constraints in 2D and 3D data sets. A comparable assignment method, which includes the correlations from the CHHC experiments, will be very useful for structure determinations of proteins in an early stage of the structure determination in the solid state.

6.3 Electronic structure of the BChl a cofactors

A major part of this thesis is focussed on the selective assignment for the three bacteriochlorophylls, B800, α B850 and β B850, to resolve their electronic ground states in the native protein binding site and to get insight into the mechanism and function of the complex. The isotropic chemical shifts in the spectrum of figure 3.1 are significantly different for all three BChl a species. After correcting the chemical shifts for the ring current effects in the B850 cofactors, the differences are minimized and the corrected chemical shifts reveal similar electronic structures for all three BChls a species. Carbons at the long axis of the BChl aromatic macrocycle in pyrrole rings A and C (figure 4.2) are changed significantly more than carbons at the short axis in pyrrole rings B and D. Clear charge polarizations along the long axis of the BChl macrocycles have been observed. This suggests that the energy levels of the ground state for the Q_y -dipole moments are changed considerably more than for the Q_x dipole moments. The electronic Q_y dipoles for both B800 and B850 species are more or less similar, which indicates that the effect of the protein environment on both the B800 and B850 is comparable and corresponds with a red shift of about 30 nm. It has been concluded that the difference between the B800 and B850 Q_y absorption band should be attributed predominantly to the excitonic coupling between the B850 cofactors, which yields splitting of the LUMO energy levels. In addition, it has been observed that the electronic changes occur rather at a molecular level than at the atomic level. This provides conclusive evidence for global electronic polarization provoked by the dielectrics of the protein environment, as opposed to local effects from e.g. polar functional groups. The electronic changes

in the BChls are attributed to mutual polarization effects between the cofactors and the protein in a nonlinear response mechanism. The LH2 protein can be considered as a medium that polarizes the whole electronic system of the BChl *a* macrocycle embedded in the protein interior, which stabilizes the complex. The global effect of the protein on the BChl *a* cofactor suggests that the protein acts as a unit and not as a sum of contributions from different important structural interactions. Although mutation studies can provide insight into the mechanism or function of a protein by changing one of the important interactions, for these studies it is difficult to survey the complete picture of the electronic structure like it is possible in NMR studies. Therefore, the application of MAS NMR dipolar correlation techniques for detailed insight into the spatial and electronic structure of the organic cofactor to reveal its mechanism is a promising biophysical approach for studying moderately sized cofactors bound to their membrane target.

Once the active binding site of a protein has been identified and analyzed, it can be very helpful to mimic its physical and chemical properties with chemical models. The development of a set of models is described in chapter 5. They are based on the high-resolution LH2 structure determined by X-ray. NMR spectroscopy reveals electronic changes of the axially coordinated Imidazole (Im) ligand and the Mg-porphyrin (Mg(II)-Por) macrocycle. Based on the chemical shift changes of the nuclei from the axial ligand and the signal intensity in the NMR spectra, it was derived that the complex exist of one partly negatively charged Mg(II)-Por and one partly positively charged axial ligand in a Mg(II)Por(δ^-)/Im(δ^+) complex with a five-coordinated Mg(II) ion. The presence of such an electronic charge distribution is confirmed in chapter 4 by the cumulative electronic changes of the carbons in the macrocycle, which indicate a partly negative charge on the BChl cofactors. The axial ligand in model studies, described in chapter 5 is more dynamic than the imidazole moiety in the histidine side chain, which coordinates to the central Mg(II) ion of BChl in photosynthetic complexes. The dynamics of the model complexes suggest that the electronic polarization of the axial ligand and the Mg(II)-Por ring of the model complexes is also in a dynamic mode. It is believed that the axial coordination and the electronic state of BChls in biological photosynthetic systems, like the light-harvesting complexes and the reaction center, is important for their high efficiency. The model systems have contributed to a clear understanding of the

electronic structure and the axial coordination to the central magnesium ion. The model complexes are useful for understanding the mechanism of the photosynthetic complexes. In addition, they are potentially also useful for the development of nature-based artificial devices for collecting solar energy.

In conclusion, the pattern labeling approach for MAS NMR experiments has appeared to be very useful for the sequence specific assignment of the LH2 protein complex, which has been largely assigned, with exception of the mobile entities, such as the loops and the C-termini. Certainly, complete assignments from N-termini to C-termini for membrane proteins will lead to a more precise determination of their structures. However, this study implies that a full protein assignment is not a prerequisite to resolve or understand its function. It is expected that focussing with solid-state NMR spectroscopy at the active site and at the cofactor(s) already provides the most important structural and electronic details, before a complete structure is known.

References

- (1) Wuthrich, K.; Wider, G.; Wagner, G.; Braun, W.; *J. Mol. Biol.* 1982, 155, 311-319.
- (2) Mulder, F. M.; Heinen, W.; van Duin, M.; Lugtenburg, J.; de Groot, H. J. M.; *J. Am. Chem. Soc.* 1998, 120, 12891-12894.
- (3) de Boer, I.; Bosman, L.; Raap, J.; Oschkinat, H.; de Groot, H. J. M.; *J. Magn. Reson.* 2002, 157, 286-291.
- (4) Lange, A.; Luca, S.; Baldus, M.; *J. Am. Chem. Soc.* 2002, 124, 9704-9705.
- (5) Sonnenberg, L.; Luca, S.; Baldus, M.; *J. Magn. Reson.* 2004, 166, 100-110.
- (6) Cornilescu, G.; Delaglio, F.; Bax, A.; *J. Biomol. NMR* 1999, 13, 289-302.
- (7) Grishaev, A.; Llinas, M.; *J. Biomol. NMR* 2004, 28, 1-10.
- (8) Herrmann, T.; Guntert, P.; Wuthrich, K.; *J. Biomol. NMR* 2002, 24, 171-189.
- (9) Grishaev, A.; Llinas, M.; *Proc. Natl. Acad. Sci. U. S. A.* 2002, 99, 10941-10941.
- (10) Herrmann, T.; Guntert, P.; Wuthrich, K.; *J. Mol. Biol.* 2002, 319, 209-227.

- (11) Mumenthaler, C.; Guntert, P.; Braun, W.; Wuthrich, K.; *J. Biomol. NMR* 1997, 10, 351-362.
- (12) Nilges, M.; Macias, M. J.; Odonoghue, S. I.; Oschkinat, H.; *J. Mol. Biol.* 1997, 269, 408-422.
- (13) Linge, J. P.; Habeck, M.; Rieping, W.; Nilges, M.; *Bioinformatics* 2003, 19, 315-316.

Summary

This thesis describes the sequence specific assignment of the photosynthetic light harvesting 2 (LH2) transmembrane protein complex and the investigation of the molecular mechanism behind this complex. The context of this work has been described in **chapter 1**. Transmembrane proteins are due to the lipophilic character of the α -helical constructed transmembrane segments not well soluble or non-soluble and are difficult to crystallize because of the interaction of the phospholipids with the lipophilic proteins. Because of this, the conventional liquid state NMR and X-ray structure determination methods are mostly not applicable and much less is known about membrane proteins than for soluble proteins. Because sooner or later almost every process in an organism is regulated through a membrane protein, it is of major importance to determine structures of these proteins and get the insight into their mechanism. Due to developments in the field of solid state NMR in the last decade, solid state NMR has the power to study extensive or uniformly labeled insoluble transmembrane proteins or protein aggregates with this technique. In this research, the LH2 complex has been chosen as a model for the application of solid state NMR to membrane proteins. The principle reason for the choice of the LH2 complex is the large abundance of this complex in the purple bacterium. Only a relatively small amount of labeled medium is needed to prepare the sample. In addition, this protein is one of the few membrane proteins with a high-resolution X-ray structure, which was helpful for the NMR studies of this thesis.

Obtaining a sequence specific NMR assignment of the amino acids in the LH2 protein is a prerequisite for the structure determination. The sequence specific assignment of the LH2 protein was realized by the labeling the protein samples with ^{13}C isotopes on a biosynthetic way according to a labeling pattern and by the use of conventional 2D correlation NMR techniques. **Chapter 2** describes the growth conditions and the composition of the growth media with [1,4- ^{13}C]-, [2,3- ^{13}C] or [1,2,3,4- ^{13}C]-succinic acid or uniformly ^{13}C , ^{15}N labeled amino acids as carbon nutrient source for the preparation of the pattern labeled 1,4-LH2, 2,3-LH2,

1,2,3,4-LH2 and the amino acid (AA)-LH2, respectively. Two-dimensional homonuclear ^{13}C - ^{13}C NMR correlation spectroscopy was used to trace the ^{13}C isotope labels from succinic acid to its destination in every individual amino acid and bacteriochlorophyll cofactor of the LH2 protein. For each amino acid and for the cofactors has been represented how the individual atoms are labeled starting from selectively labeled succinic acid, leading to an effective increase of the resolution. Because of the predominantly α -helical structural elements in the LH2 complex there is a narrow distribution and strong overlap of the ^{13}C correlation signals of the backbone carbons. The preparation of various pattern labeled samples is essential, because in this way the crowded spectra can be subdivided in to multiple datasets that are less crowded. After determination of the labeling pattern, these LH2 samples are used to obtain a sequence specific chemical shift assignment in chapter 3.

For a sequence specific assignment, the resonances of isotopically labeled nuclei from adjacent residues should be correlated. This procedure has been described in **chapter 3**. Correlations between nuclei were obtained from 2D homonuclear ^{13}C - ^{13}C and heteronuclear ^{13}C - ^{15}N correlation experiments. The characteristic correlation patterns in the ^{13}C - ^{13}C correlation spectra enable the identification of the ^{13}C responses of each amino acid type in the protein. By using the band selective heteronuclear ^{13}C - ^{15}N polarization technique, two types of spectra were obtained, the NCACX and the NCOACX correlation spectrum. In the NCACX correlation spectrum the ^{15}N nucleus correlates with carbon atoms of the same residue. The NCACX together with the ^{13}C - ^{13}C correlation spectra provides a complete set of resonances for individual amino acids. In the NCOACX spectrum, the ^{15}N nucleus correlates with carbon atoms from a previous residue of the protein sequence, determining the identification of the position of the residue in the protein sequence, which is the sequence specific assignment.

This chapter also describes how the sequence specific assignment could be obtained on the basis of ^{13}C - ^{13}C correlation experiments with a spin diffusion period of 500 ms. Due to the reduction of the number of labels in the pattern labeled LH2 samples it is possible to transfer the magnetization over larger distances than in case of the uniformly labeled protein sample. Because of this, not only intra-residue but also inter-residue correlations have been observed. These correlation

experiments have strongly contributed to the sequence specific assignment of 76 out of 94 residues in the LH2 complex. Moreover, also protein-cofactor correlations have been observed in these spectra, providing important information about the electronic structure of the active site. 18 Amino acids have not been observed because they are located in a mobile part of the protein and do not have a homogeneous structure.

Chapter 4 concentrates on the bacteriochlorophyll cofactors, which are of major importance for the function and mechanism of the LH2 protein. Bacteriochlorophyll cofactors in the LH2 complex absorb at longer wavelength than dissolved in acetone. The red shift of the bacteriochlorophylls in the LH2 complex is potentially due to different electronic structures of the ground states. The LH2 complex has two types of bacteriochlorophylls for each monomeric LH2, one B800 and two B850 (α B850 and β B850), denoted to the absorption maximum in the UV-vis spectrum. Solid state NMR spectroscopy has been applied to resolve the electronic ground states of the bacteriochlorophyll ground states.

The solid state NMR signals of the bacteriochlorophylls in the 2D correlation spectra of the uniformly [^{13}C , ^{15}N]-labeled LH2 sample have been assigned selectively to B800, α B850 and β B850. The chemical shift differences, obtained by comparing these signals with corresponding signals from isolated bacteriochlorophyll dissolved in acetone- d_6 , provide insight into the electronic structure of the bacteriochlorophylls. Solid state NMR spectroscopy has shown that the electronic structure of the cofactor is due to global interactions between the protein and the cofactor, rather than local interactions between protein and cofactor. The electronic ground states of B800 and B850 cofactors are nearly equal, but differ significantly from the electronic ground state of bacteriochlorophyll in acetone solution. The protein environment plays an important role for the position of the absorption maximum in the UV-vis spectrum. The difference of the absorption wavelength between B800 and B850 is mainly a result of the mutual interactions between B850 cofactors.

Also of concern for the electronic structure of the cofactor is the specific coordination from histidine residues in the protein to magnesium ion in the center of the cofactors. **Chapter 5** describes the advanced chemical models that have been used to study the active site of the complex, in particular for a detailed insight into

the coordination bond between the histidine residue and the magnesium ion. ^1H NMR experiments have shown that the magnesium ion in the porphyrin of the model complex has the same five-coordination, in correspondence with the natural protein complex. ^1H and ^{13}C NMR spectra show large chemical shift changes between 0 and 5.5 ppm for the coordinated ligand. These chemical shift changes are predominantly induced by the ring current effects of the porphyrin ring. A chemical shift change of 25 ppm is observed for the nucleus of the ^{15}N atom of the axial ligand that coordinates to the magnesium ion. This shift shows a partially positively charged axial ligand. This charge is compensated by a partially negative charge on the porphyrin ring. These results are well in line with a partial negative charge on the bacteriochlorophyll cofactors, which is observed with solid state NMR and described in chapter 4. The chemical models have contributed to a detailed insight into the electronic structure of the cofactors in the LH2 complex.

A general discussion and future perspectives of the research are elaborated in **chapter 6**. It is anticipated that the method described in this thesis in future will be applied on other membrane proteins as well and will lead to the determination of membrane protein structures. The possibility to resolve 3D membrane protein structures and protein-cofactor interactions means that solid state NMR will have an important role in the study of membrane proteins.

Samenvatting

Dit proefschrift beschrijft de sequentiespecifieke toekenning van het fotosynthetisch light-harvesting 2 (LH2) transmembraaneiwitcomplex en het onderzoek naar het mechanisme van dit complex. De context van dit werk wordt beschreven in **hoofdstuk 1**. Transmembraaneiwitten zijn vanwege het lipofiele karakter van de α -helices moeilijk of niet oplosbaar en, door de fosfolipiden die met de lipofiele eiwitcomplexen interactie hebben, zeer moeilijk te kristalliseren. Hierdoor zijn de conventionele vloeistof-NMR- en kristallografie methoden voor het bepalen van de structuur van deze eiwitten meestal niet toepasbaar en is er veel minder bekend over membraaneiwitten dan over oplosbare eiwitten. Omdat vrijwel elk proces in een organisme vroeger of later via één of meerdere membraaneiwitten wordt gereguleerd, is het bepalen van de structuur en inzicht in het mechanisme van deze eiwitten van groot belang. Door de ontwikkelingen op het gebied van vaste-stof-NMR in het laatste decennium is het mogelijk om met deze techniek ook uniform gelabelde, onoplosbare transmembraaneiwitten en eiwitaggregaten te bestuderen. Als model voor de toepassing van vaste-stof-NMR-spectroscopie op membraaneiwitten is in dit onderzoek het LH2-complex gekozen. De voornaamste reden voor de keuze van het LH2-eiwit ligt in het feit dat dit eiwit in grote hoeveelheden kan worden geïsoleerd. Slechts weinig medium is nodig is voor het maken van een preparaat. Ook is het LH2-complex één van de weinige transmembraaneiwitten, waarvan een kristalstructuur met röntgendiffractie is vastgesteld.

Het verkrijgen van een sequentiespecifieke NMR-signaaltoekenning voor de aminozuren in het LH2-eiwit is een eerste vereiste voor de structuurbepaling. De sequentiespecifieke toekenning van het LH2-eiwit is tot stand gekomen door de eiwitpreparaten op biosynthetische wijze volgens een bepaald patroon met ^{13}C -isotopen te labelen en bestaande 2D-NMR-correlatietechnieken te gebruiken. **Hoofdstuk 2** beschrijft de groeiomstandigheden en de samenstelling van de groei-media met [1,4- ^{13}C]-, [2,3- ^{13}C]-, [1,2,3,4- ^{13}C]-succinaat of uniform ^{13}C , ^{15}N -gelabelde aminozuren als koolstofbron en de groeiomstandigheden voor het prepareren van respectievelijk het patroongelabelde 1,4-LH2, 2,3-LH2, 1,2,3,4-LH2 en het aminozuren(AA)-LH2 preparaat. Tweedimensionale homonucleaire ^{13}C - ^{13}C -NMR

correlatiespectroscopie werd gebruikt om de ^{13}C -isotooplabels van het succinaat naar de bestemming in elk individueel aminozuur en bacteriochlorofyl co-factor van het LH2-eiwit te traceren. Voor elk aminozuur en voor de co-factoren is een overzicht gepresenteerd dat weergeeft hoe de individuele atomen gelabeld zijn, uitgaande van het selectief gelabeld succinaat. Ook wordt aangetoond dat de spectrale dichtheid van de patroongelabelde LH2-preparaten ten opzichte van het correlatiespectrum van het uniform gelabelde LH2-preparaat is afgenomen, terwijl de resolutie in de spectra is toegenomen. Omdat het LH2-complex overwegend is opgebouwd uit α -helices, is er nauwelijks spreiding en een sterke overlap van de NMR-signalen van de ^{13}C -kernen in de hoofdketen van het eiwit. Het labelen van LH2-complexen volgens diverse labelpatronen was noodzakelijk, omdat op deze wijze de “drukke” correlatiespectra van het uniform gelabelde LH2-complex konden worden opgedeeld in meerdere “minder drukke” spectra. Voorts werden deze LH2-preparaten gebruikt voor het verkrijgen van een sequentiespecifieke toekenning van de chemische verschuivingen in hoofdstuk 3.

Voor een sequentiespecifieke toekenning is het noodzakelijk de resonantiesignalen van ^{13}C - en ^{15}N -isotoopgelabelde kernen van aangrenzende residuen met elkaar te correleren. Deze procedure wordt beschreven in **hoofdstuk 3**. Correlaties tussen kernen werden verkregen uit 2D-homonucleaire ^{13}C - ^{13}C - en heteronucleaire ^{13}C - ^{15}N -correlatiespectra. De karakteristieke correlatiepatronen in de ^{13}C - ^{13}C -homonucleaire spectra maken het mogelijk de ^{13}C -signalen van elk type aminozuur in het eiwit te identificeren. Met de selectieve heteronucleaire ^{13}C - ^{15}N -polarisatietechniek werden 2 typen spectra verkregen; het NCACX- en het NCOCOCX-correlatiespectrum. In het NCACX-correlatiespectrum correleert een ^{15}N -kern met de ^{13}C -kernen van hetzelfde residu en in het NCOCOCX-correlatiespectrum correleert een ^{15}N kern met ^{13}C -kernen van een vorig residu in de eiwitsequentie. De NCACX-spectra geven samen met de ^{13}C - ^{13}C -homonucleaire spectra een complete set van resonanties voor individuele aminozuren. Het NCOCOCX-spectrum laat interresidu correlaties zien. In het NCOCOCX-spectrum correleert de ^{15}N -kern met ^{13}C -kernen van een vorig residue, wat de identificatie van de positie van het residu in de eiwitsequentie bepaalt.

Tevens wordt in dit hoofdstuk beschreven hoe een sequentiespecifieke toekenning verkregen kon worden op basis van ^{13}C - ^{13}C -correlatieexperimenten met een spindiffusieperiode van 500 ms. Door een reductie van het aantal labels in de

volgens een patroon gelabelde LH2-preparaten is het mogelijk om de magnetisatie over grotere afstand over te dragen dan in het geval van een uniform gelabeld LH2-preparaat. Daardoor worden niet alleen intraresidu, maar ook interresidu ^{13}C - ^{13}C -correlaties waargenomen. Deze correlatie-experimenten hebben in belangrijke mate bijgedragen aan de sequentiespecifieke toekenning voor 76 van de 94 residuen in het LH2-eiwitcomplex. Bovendien worden in deze spectra ook interacties tussen het eiwit en de co-factoren waargenomen. Deze interacties geven belangrijke informatie over de elektronische structuur van het actieve centrum. Een 18-tal aminozuren werd niet geobserveerd, omdat de betreffende aminozuren in een mobiel gedeelte van het eiwit zitten en geen homogene eiwitstructuur hebben.

In **hoofdstuk 4** wordt verder ingegaan op de bacteriochlorofyl co-factoren die van groot belang zijn voor de functie en het mechanisme van het LH2-eiwit. De bacteriochlorofyl co-factoren in het LH2-complex absorberen bij langere golflengten dan wanneer deze zijn opgelost in aceton. Deze roodverschuiving van de bacteriochlorofyl co-factoren in het LH2-complex is mogelijk te wijten aan veranderingen van de elektronische structuur van de grondtoestanden. Het LH2-complex heeft twee typen bacteriochlorofyllen per LH2-monomeer; één B800 en twee B850 (α B850 and β B850), welke genoemd zijn naar hun absorptiemaximum in het optische spectrum. Vaste-stof-NMR-spectroscopie wordt toegepast om de elektronische grondtoestanden van de bacteriochlorofyl-cofactoren vast te stellen.

De vaste-stof-NMR-signalen van de bacteriochlorofyl-co-factoren in het 2D-homonucleaire ^{13}C - ^{13}C -correlatiespectrum van het uniform [^{13}C , ^{15}N]-gelabelde LH2-preparaat worden selectief aan B800, α B850 en β B850 toegekend. De verschillen in chemische verschuiving, verkregen door de NMR-signalen te vergelijken met overeenkomstige signalen van geïsoleerd bacteriochlorofyl opgelost in aceton- d_6 , werden gebruikt om inzicht te krijgen in de elektronische structuur van de bacteriochlorofyllen. Vaste-stof-NMR-spectroscopie heeft aangetoond dat de elektronische structuur van de co-factoren het gevolg is van de algehele interactie tussen eiwit en de co-factor, en niet primair bepaald wordt door locale interacties van de co-factor met het eiwit. De elektronische grondtoestanden van de B800 en de B850-co-factoren zijn nagenoeg identiek, maar verschillen significant met bacteriochlorofyl in oplossing van aceton. De eiwitomgeving speelt een belangrijke rol voor de positie van het absorptiemaximum. Het verschil in absorptie golflengte

tussen B800 en B850 is met name het gevolg van onderlinge interacties tussen de B850-co-factoren.

Mede van belang voor de elektronische structuur van de co-factor is de specifieke coördinatie van de histidineresiduen in het eiwit naar het magnesium-ion in het centrum van de bacteriochlorofyl-co-factoren. In **hoofdstuk 5** worden de chemische modellen beschreven die gebruikt zijn om het actieve centrum van het complex, met name de coördinatieband tussen het axiale histidineresidu en het magnesium-ion nader te bestuderen. ^1H -NMR-experimenten hebben aangetoond dat de porfyriene in de modelverbinding dezelfde unieke 5-coördinatie rond het magnesium-ion heeft als in het natuurlijke eiwitcomplex. ^1H - en ^{13}C -NMR laten grote verschuivingen tussen 0 en 5.5 ppm voor het gecoördineerde ligand zien. Deze verschuivingen zijn met name het gevolg van het ringstroomeffect in de porfyriering. Een chemische verschuiving van 25 ppm is waargenomen voor de kern van het coördinerende ^{15}N -atoom van het axiale ligand. Deze verschuiving laat zien dat het axiale ligand gedeeltelijk positief geladen is. Deze lading wordt gecompenseerd door een gedeeltelijke negatieve lading op de porfyriering. Deze resultaten zijn in overeenstemming met de gedeeltelijke negatieve lading van de co-factoren, welke is waargenomen met behulp van vaste-stof-NMR-spectroscopie en beschreven in hoofdstuk 4. De chemische modellen hebben bijgedragen aan een gedetailleerd inzicht in de elektronische structuur van de co-factoren in het LH2-complex.

Een algemene discussie met de perspectieven van het onderzoek wordt weergegeven in **hoofdstuk 6**. Het ligt in de lijn der verwachting dat de methode beschreven in dit proefschrift in de toekomst ook wordt toegepast op andere transmembraaneiwwitten en zal leiden tot de bepaling van membraaneiwitstructuren. De mogelijkheid om 3D-structuren van membraaneiwwitten en interacties tussen eiwit en co-factoren te bepalen, betekent dat er een belangrijke rol is weggelegd voor vaste-stof-NMR-spectroscopie in de studie van de membraaneiwwitten.

List of abbreviations

Å	Angström (one Å = 0.1 nm)
1D	One-dimensional
2D	Two-dimensional
B800, B850	Bacteriochlorophylls with Q _y -band absorption at 800 and 850 nm.
(B)Chl	(Bacterio)chlorophyll
CP	Cross Polarization
δ	chemical shift in solution NMR
Δδ	difference between chemical shifts
DFT	Density Functional Theory
GIAO	Gauge Invariant Atomic Orbital
H/His	Histidine
HMBC	Heteronuclear Multiple Bond Correlation
HOMO	Highest Occupied Molecular Orbital
Im	Imidazole
IR	Infra Red
kDa	kiloDalton
LDAO	Lauryldimethylamine <i>N</i> -oxide (<i>N,N</i> -dimethyldodecylamine- <i>N</i> -oxide)
LH1	Light-Harvesting 1
LH2	Light-Harvesting 2
LUMO	Lowest Unoccupied Molecular Orbital
MAS	Magic Angle Spinning
1-MeIm	1-methylimidazole
Mg(II)-Por	Mg(II)-porphyrin macrocycle
NICS	Nuclear Independent Chemical Shift
NMR	Nuclear Magnetic Resonance
OD	Optical Density
ODV	Optical Density per Volume
OEP	Octaethylporphyrin
PDS	Proton Driven Spin Diffusion
<i>R.</i>	<i>Rhodospseudomonas</i>

<i>Rb.</i>	<i>Rhodobacter</i>
RC	Reaction Center
r.f.	Radio Frequency
RFDR	Radio Frequency Dipolar Recoupling
r_m	molar ratio
$\Delta\sigma$	chemical shift difference in the solid state
σ^l	chemical shift in acetone- d_6 solution
σ^s	isotropic chemical shift in solid state NMR
σ^R	ring current shift induced by bacteriochlorophylls
$\Delta\tilde{\sigma}$	chemical shift difference due to electronic changes
SSNMR	Solid State NMR
THF	tetrahydrofuran
TPPM	Two Pulse-Phase Modulation
TRIS	tris-(hydroxymethyl)-aminomethane
UV-vis	ultraviolet visible

Publications

1. Van Gammeren, A. J., Hulsbergen, F. B., Erkelens, C. & de Groot, H. J. M. (2004) *Journal of Biological Inorganic Chemistry* 9, 109-117.
Synthetic analogues of the histidine-chlorophyll complex: a NMR study to mimic structural features of the photosynthetic reaction center and the light-harvesting complex.
2. Van Gammeren, A. J., Hulsbergen, F. B., Hollander, J. G. & de Groot, H. J. M., (2004) *Journal of Biomolecular NMR* 30, 267-274.
Biosynthetic site-specific ^{13}C labeling of the light-harvesting 2 protein complex as a model for solid state NMR spectroscopic studies on transmembrane proteins.
3. Van Gammeren, A. J., Buda F., Hulsbergen, F. B., Kiihne S., Hollander, J. G., Egorova-Zachernyuk, T. A., Fraser, N. J., Cogdell, R. J. & de Groot, H. J. M., (2005) *Journal of the American Chemical Society*. In press.
Selective chemical shift assignment of B800 and B850 bacteriochlorophylls in uniformly [^{13}C , ^{15}N] labeled light-harvesting 2 complexes by solid state NMR spectroscopy.
4. Van Gammeren, A. J., Hulsbergen, F. B., Hollander, J. G. & de Groot, H. J. M., (2005) *Journal of Biomolecular NMR*. In press.
Backbone and side-chain ^{13}C and ^{15}N resonance assignments of the intrinsic transmembrane light-harvesting 2 protein complex by solid state Magic Angle Spinning NMR.

Curriculum vitae

Na in 1996 het VWO diploma aan het Van Lodenstein College te Amersfoort te hebben behaald, is de auteur in september van hetzelfde jaar begonnen met de studie Scheikunde aan de Universiteit Leiden. In 1997 werd het propedeutisch examen gehaald. In het kader van zijn afstudeeronderzoek heeft hij van februari 1999 tot en met januari 2001, onderzoek verricht naar de organische synthese van isotoopverrijkte aminozuren bij de werkgroep Bio-organische Fotochemie onder leiding van prof. dr. J. Lugtenburg. Tijdens deze periode heeft hij een zomerstage bij de afdeling Biophysical Chemistry van het Occidental College in Amerika (Californië) onder leiding van prof. dr. P. Dea gedaan. Daar werd onderzoek verricht aan de incorporatie van chloorpromazine-hydrochloride in fosfatidylcholine membranen. In februari 2001 werd het doctoraaldiploma behaald.

In maart 2001 werd bij de werkgroep Biofysische Organische Chemie onder leiding van prof. dr. H. J. M. de Groot begonnen met het promotieonderzoek "membrane protein studies with magic angle spinning NMR". De resultaten van dit onderzoek, zoals beschreven in dit proefschrift, werden gepresenteerd aan de hand van posterbijdragen op diverse nationale en internationale bijeenkomsten en congressen, te weten het *XXth International Conference on Magnetic Resonance in Biological Systems (ICMRBS)* in Toronto (Canada, 2002), het *European Experimental NMR/Ampère conference* in Lille (Frankrijk, 2004), het *International Society of Magnetic Resonance (ISMAR) congress* in Jacksonville (Florida, U. S. A., 2004), enkele CW-studiegroepen te Lunteren en tijdens de NMR Discussion Group dagen te Vlaardingen, Groningen en Utrecht. Daarnaast werden door de auteur de resultaten door middel van enkele voordrachten gepresenteerd, te weten het *BIOMAC AIO-OIO Symposium* te Leiden (2002), het *International Festsymposium zur Inbetriebnahme des 900 MHz-Spektrometers am Forschungsinstitut für Molekulare Pharmacologie* te Berlijn (Duitsland, 2004) en de SON studiegroep *Spectroscopie en Theorie* te Lunteren (2005).

Nawoord

Met het verschijnen van dit proefschrift is een einde gekomen aan mijn Leidse studieperiode. En dan denk ik ...hora ruit, tempus fluit... het uur snelt heen, de tijd vergaat. Een beetje schrik ik van de snelheid waarmee de tijd vervlogen is. Een goed moment om even stil te staan en terug te blikken in het leven dat steeds sneller lijkt te gaan.

Hoewel het schrijven van een proefschrift een overwegend solitaire bezigheid is, was het tot stand komen slechts mogelijk door samenwerking met en steun van anderen. Het proefschrift zou niet af zijn zonder deze laatste bladzijde die ik graag wil gebruiken om hen te noemen die, hetzij wetenschappelijk hetzij sociaal, een grote of kleine bijdrage hebben geleverd.

De isotoopverrijkte eiwitpreparaten voor dit onderzoek werden door mij geprepareerd bij de afdeling Biofysica in het Huygens laboratorium. Dankzij Dré de Wit voelde ik mij daar snel thuis en kon een snelle start maken. De vaste-stof-NMR-spectra van de isotoopverrijkte preparaten werden veelal in samenwerking met Johan Hollander gemeten. Johan, ik ben je daarvoor zeer erkentelijk. Niet zelden waren er problemen met de probe of met de meting en altijd weer stond je klaar om deze op te lossen. Ook Fons Lefeber en Kees Erkelens waren altijd paraat en hebben naast Johan voor veel levendigheid en gezelligheid op de NMR-afdeling gezorgd. Van groot belang waren ook de diepgaande inhoudelijke discussies met Francesco Buda en Suzanne Kiihne.

Een onvergetelijke tijd dank ik aan alle AIO's van de SSNMR en de BOF groep. Met Ido de Boer heb ik, naast de inhoudelijk wetenschappelijke discussies over onze onderzoeken, ook veel gesprekken gevoerd over de meest uiteenlopende onderwerpen. I will remember the pleasant time I had with Prashant Agrawal. Prashant, the time we spent together was wonderful. I will always remember your company, your culinary Indian recipes and your readiness to help. I want to thank Anna Karawajczyk for the nice time in our "denkcel" and Dorota Bartusik for the work she did with me in our laboratory during six months.

Het belang van mentale steun van familieleden en vrienden is voor de totstandkoming van dit werk niet te onderschatten. In het bijzonder gaat mijn dank uit naar mijn ouders voor hun liefde, vertrouwen en steun. Zij hebben mij altijd bijgestaan en bemoedigd in mijn keuzes. Door het klussen in ons nieuwe huis waren de weekends van het afgelopen jaar weliswaar erg druk. Toch was het fysieke werk, waarbij mijn schoonouders altijd klaarstonden, een welkome afleiding tijdens het drukke schrijfproces van dit proefschrift.

Tenslotte, maar eigenlijk in de eerste plaats ben jij, Mirjam, degene die telkens weer een onontbeerlijke steun was en mij daardoor het meest heeft bijgestaan. Niet alleen je betrokkenheid, ook je realisme en scherpe geest waren in tijden dat het wel eens tegen zat een bemoediging om weer door te zetten om dit proefschrift te realiseren.

Leiden, april 2005

Adriaan van Gammeren