

Chapter 3

NMR investigations of bacteriorhodopsin

3.1 NMR strategies for structural investigation of membrane proteins

Membrane proteins comprise approximately one third of the genome of an organism. They are responsible for most membrane functions, serving as receptors, regulating transport across the membrane or working as enzymes for example. In spite of their biological importance less than 25 high-resolution structures of membrane proteins have been published today by employing both electron and X-ray crystallography (Kühlbrandt and Gouaux, 1999). The progress in structure investigations is relatively slow, especially when compared to the rapidly growing number of structures for water-soluble proteins. Progress for membrane proteins has been hampered by their low abundance (Grishammer and Tate, 1995) and the instability outside the membrane. However, the major difficulty for diffraction methods is growing well-ordered two- or three-dimensional crystals (Michel, 1991).

This last step is not needed for NMR structure investigations, which makes NMR an attractive alternative to the more established diffraction methods. Solubilized proteins can be investigated with the same high-resolution liquid state NMR methods applied to the study of water-soluble proteins. The maximum tractable protein size for such studies is however reached very quickly, since the bound detergent molecules increase the effective molecular weight. The resulting slow rotational tumbling of the protein-detergent complex leads to the same difficulties as for large water-soluble proteins (see chapter 1.3). Backbone assignments of membrane proteins have been obtained so far only for samples with less than 75 amino acids or two α -helical segments (Henry and Sykes, 1994; Opella, 1994; Opella, 1999; Pervushin, 1994; Williams, 1996). The size problem can be reduced by the use of organic solvents. Assignments for a 110 residue protein in a mixture of chloroform/methanol/water have been obtained. The NMR data suggest a structure with four α -helices (Schwaiger, 1998), but as in the studies in micelles cited above there were not enough experimental data to define the tertiary structure unambiguously from the liquid state NMR data. This might be due to slow chemical exchange processes or due to multiple tertiary conformations, making the observation of long-range NOEs between neighboring helices impossible.

From a biological point of view, the medium stabilizing the protein structure should resemble the native membrane environment as close as possible. Detergent micelles are certainly to be preferred over organic solvents with this respect, but it is to question whether the small diameter of the micelles induce some artificial effects. An alternative medium is phospholipid bilayers, which have a nearly planar surface. Lipid bilayers are essentially solid from a NMR point of view and membrane proteins in lipid bilayers are therefore usually subject of solid state NMR investigations (Griffin, 1998; Opella, 1994; Opella, 1997; Smith , 1996).

The major difficulty of solid state in comparison liquid state NMR investigations is the increase in linewidth of the signals due to the dipolar (and quadrupolar for spins $> 1/2$) interaction of the spins. The problem can be overcome partially by fast mechanical rotation of the whole sample, leading to an averaging of the interaction around the axis of rotation. If the rotation axis is chosen to be at the magic angle $\theta=54.74$ with respect to the external magnetic field, the angular term $3\cos^2\theta-1$ characteristic for the secular contribution of the dipolar interaction vanishes. This leads to high resolution solid state spectra in cases where the rotation frequency is faster than the characteristic linewidth of the signals (Stejskal, 1994). With current technology, it should be possible to resolve and assign the carbon signals of completely isotope labelled proteins up to 30 amino acids (Tycko, 1996). Several research groups currently work out the development of experimental strategies for resonance assignment and distance determination in small, uniformly labelled proteins.

The resonance frequency of a spin in solid state depends on its orientation with respect to the external magnetic field. This leads to so-called “powder pattern” lineshapes, if the molecules are oriented at random. In samples with a higher macroscopic order, microscopic structural parameters can be derived from the observed NMR signals. Alignment of membrane protein samples can be achieved by mechanical orientation of lipid bilayers or by using the magnetic susceptibility anisotropy of phospholipid bilayers (Sanders , 1994; Sanders and Landis, 1995) containing the protein. A selectively labelled membrane protein fragment prepared in such phospholipid bilayers has been studied using heteronuclear liquid crystal NMR techniques (Losonczi and Prestegard, 1998).

The information obtained from macroscopically oriented samples is complementary to that obtained from samples in small micelles. As has been mentioned above, experiments in micelles often lack information about the three dimensional arrangement of identified secondary structure elements. This information can be obtained on oriented samples and an approach which combines solid state and liquid state experiments has been applied successfully to segments of one and two α -helices. The secondary structure has been determined on the basis of short internuclear distance based on proton proton NOEs measured in micelles. The arrangement of the secondary structure elements with respect to the membrane normal and their relative orientation has been derived from measurements of angular parameters in oriented bilayer samples (Opella , 1994; Opella , 1999).

Bacteriorhodopsin is a quite exceptional case of an integral membrane protein. It is of moderate size and can be overexpressed and purified to yield large amounts of protein. As for other biophysical methods, it served as a model compound for many NMR strategies applied to membrane proteins.

Historically the first approach was magic angle spinning solid state NMR on randomly oriented purple membrane. In the first experiments, the protein was selectively labelled in the ϵ -[^{15}N] position of all lysine side chains (Harbison , 1983). Due to the different chemical environment of the retinal Schiff-base ^{15}N compared to the rest of the lysine side chain ^{15}N and the backbone ^{15}N nuclei, the Schiff-base signal is resolved in a one dimensional magic angle spinning spectrum. The chemical shift of the ^{15}N Schiff base resonance has been shown later to obey a direct proportionality to the absorption maximum of the retinal Schiff base compound, which in turn depends on the counterion strength (de Groot , 1989; de Groot , 1990; Hu , 1994). Bacteriorhodopsin with selective single ^{13}C labels in the retinal have been prepared by reconstitution of bacterioopsin with synthetically labelled retinal (Harbison , 1984a). The approach lead to the identification of the two isomeric forms of the retinal in dark-adapted bacteriorhodopsin (Harbison , 1985; Harbison , 1984b). In this way, the elements of the chemical shift tensors of all carbon atoms in the retinal have been determined (Smith , 1989b). The approach was extended to the study of different

photointermediates, which were enriched by illumination combined with variations of the pH, salt concentration and temperature of the samples (Farrar , 1993; Hu , 1998; Hu , 1997; Lakshmi , 1994; McDermott , 1994; Smith , 1989a). The structural interpretation of the data in all of the above-cited investigations is based on semi-empirical chemical shift arguments. The chemical shift of the same spin in model compounds is compared to the corresponding shift in the protein. Direct distance measurements have been obtained for single pairs of spins by rotational resonance (Creuzet , 1991; McDermott , 1994; Thompson , 1992) or radio frequency-driven dipolar recoupling (Griffin, 1998).

Angular constraints for the retinal orientation with respect to the membrane normal in bacteriorhodopsin have been derived from deuterium NMR on oriented purple membrane. In this approach single methyl groups in the retinal are deuterated. For a spin 1 nucleus (deuterium) the quadrupolar interaction (i.e. the interaction between the magnetic quadrupole moment of the nucleus with the electric field gradient at the site of the nucleus) is of similar or bigger size than the dipolar interaction. Deuterium spectra of macroscopically aligned samples show a line splitting, which depends on the angle Θ between the quadrupolar tensor and the external field. This approach has been used to determine the orientation of the methyl groups 18, 19, 20 (Ulrich , 1992; Ulrich , 1994) and one of the methyl groups connected to C1 (Moltke , 1998) in the all-trans ground state and of the methyl group 19 in the M intermediate (Ulrich , 1995) by site-specific deuterium labelling of the corresponding methyl groups (see Fig. 3.2 for the numbering of the carbon atoms). The accuracy and the sensitivity of the method can be increased considerably by magic angle spinning of the oriented samples (Glaubitx , 1999).

Liquid-state NMR experiments have been performed on bacteriorhodopsin solubilized in different detergent micelles and in an organic solvent. Arseniev et al. have performed ^1H - ^{15}N studies in methanol/chloroform (1:1), 0.1M $^2\text{HCO}_2\text{NH}_4$ and SDS micelles of the complete protein and of different fragments (Barsukov , 1990; Barsukov , 1992; Efremov , 1999; Grabchuk , 1996; Lomize , 1992; Nolde , 1997; Orekhov , 1992; Orekhov , 1994; Pashkov , 1996; Pervushin and Arseniev, 1992; Pervushin , 1994; Sobol , 1992). So far, this approach did not provide any constraints for the three-dimensional packing of the α -helices in bacteriorhodopsin. Different

detergents for the solubilization of bacteriorhodopsin have been tested by Seigneuret et al. (Seigneuret, 1991b). A limited number of assignments have been obtained by labelling the protein specifically with methyl- ^{13}C (Seigneuret, 1991a) or carbonyl- ^{13}C (Seigneuret and Kainosho, 1993) methionine.

3.2 Labelling and assignment strategy used in this thesis

Bacteriorhodopsin solubilized in dodecyl-maltoside micelles forms a particle with a rotational correlation time of approximately 40 ns at 45°C (see Appendix C). In general, there are two different labelling concepts for solution-state NMR on large proteins: (1) ^{13}C , ^{15}N and ^2H enrichment and (2) uniform deuteration of the protein with selective protonation of individual units in the system (LeMaster, 1994). For the present study, the second approach seemed to be more promising for several reasons. As pointed out in chapter 1.3, a successful backbone assignment is most likely to be possible with a sample in which all non-exchangeable protons are replaced by deuterium. The relative small dispersion of the signals expected in a ^1H - ^{15}N heteronuclear spectrum for proteins of mainly α -helical conformation together with the large expected linewidths of the signals might lead to major difficulties in the assignment process. After assignments have been made, the main structural constraints derived from such a sample would be NH_i - NH_{i+1} and NH_i - NH_{i+3} NOESY cross-peaks expected for α -helices. In view of the fact that a low-resolution structural model of bacteriorhodopsin obtained by electron-crystallography was available at the start of the project already (Henderson, 1990), which was refined subsequently to a resolution of 3.5 Å (Grigorieff, 1996), identification of secondary structure topology did not seem the first goal.

The resolution of the electron diffraction data is not high enough to resolve the two closely overlapping retinal conformations (all-trans, 15-anti and 13-cis, 15-syn). We therefore concentrated on obtaining direct structural information of the chromophore by the second labelling approach mentioned above. The structural information available was used to select the amino acids to be selectively protonated in an otherwise deuterated protein.

The labelling pattern was achieved by growing cells in a fully deuterated medium complemented with ^1H -retinal, ^1H -threonine, ^1H -tryptophan, ^1H -methionine, ^1H -lysine

or ^1H tyrosine (Patzelt, 1997). Spectra from samples with the following combinations of protonated residues were analyzed in detail:

1. Retinal (d-DM)
2. Tyrosine (d-DM)
3. Tryptophan, Tyrosine (h-DM)
4. Tryptophan, Threonine (h-DM)
5. Retinal, Tryptophan (h-DM, d-DM)
6. Retinal, Threonine (d-DM)
7. Retinal, Tryptophan, Tyrosine (d-DM)
8. Retinal, Tryptophan, Threonine (d-DM)
9. Retinal, Tryptophan, Methionine (d-DM)
10. Retinal, Lysine, Methionine (d-DM)
11. Retinal, Lysine, Threonine (d-DM)
12. Retinal, Tyrosine (h-DM)

In parentheses is indicated, whether the protein has been solubilized in fully protonated dodecyl-maltoside (h-DM) or in dodecyl-maltoside with a deuterated dodecyl moiety (d-DM).

3.3 Practical aspects of NMR measurements and data processing

All NOESY spectra were recorded in a time range of about 2 years on two different NMR spectrometers operating at a ^1H frequency of 750 MHz, except for one spectrum recorded at 800 MHz. The measurement time for one spectrum at a certain mixing time was between 1.5 and 2.5 days. Spectra with different mixing times on one particular sample were typically started automatically one after the other without interference of the user in between two spectra, leading to a total of about 7 days of data acquisition. In all spectra, the protein signals, which are the signals of interest, are smaller by a factor of about 100-500 compared to the sugar signals of the detergent.

The long measurement times together with the large unwanted signals of the detergent lead to a number of problems:

1. Artifact free Fourier transformation requires proper digitization of all signals; thus the strongest signal determines the receiver gain. If very small signals are recorded at the same time, they are not optimally digitized in the A/D converter. On Bruker spectrometers the optimal signal-to-noise ratio is normally achieved, if the

receiver gain is 128 or larger. For the NOESY spectra, the receiver gain was typically between 16 (for samples with protonated detergent tail) and 64 (for samples with deuterated detergent tail).

2. Free induction decays should be recorded from $t = 0$, e.g. the acquisition should start immediately after the last 90° pulse of the NOESY sequence. Since the transmitter and receiver channel use the same coil and have to be separated electronically to avoid damage of the sensitive preamplifier by the high power output of the transmitter, acquisition necessarily is delayed after the last pulse and the intensity of the first points is influenced by the opening of the receiver gate. This leads to a shift of the signal in the time domain and small intensity errors of the first points. Intensity errors in the first points of the FID lead to slowly varying undulations of the baseline in the spectrum (Hoch and Stern, 1996).
3. Temperature instabilities and the drift of the static magnetic field lead generally to a decrease of the quality of the spectra recorded later in a series. This is usually accompanied by a drop of the lock signal. The effect is pronounced in some of the spectra, making them practically useless for a quantitative analysis of peak intensities.
4. A general source of systematic artifacts in NOESY spectra is the evolution of zero quantum coherence of J-coupled spins during the mixing time (Neuhaus and Williamson, 1989). The corresponding coherence transfer pathway can't be suppressed by phase cycling or pulsed field gradients, because it involves the same coherence orders as the pathway of the wanted signal. When there is no NOE between the coupled spins, the zero-quantum peak has a dispersive (antiphase) peak-shape with respect to the in-phase (absorptive) NOEs. The addition of a NOESY cross-peaks leads to a distorted tilted peak (Bodenhausen, 1984). For the slow tumbling rates, the antiphase components of the artifacts partially cancel due to overlap of the positive and negative parts of the cross-peak. This makes the effect so small that it does not distort the protein NOEs considerably. For the detergent signals this is different. The cross-peaks between e.g. H1 and H2 of the sugar-head group have pronounced antiphase characteristic. Because of t_1 noise ridges, the spectral region of detergent cross-peaks can not be used for the analysis of protein signals. Therefore the additional perturbations do not pose a practical problem for regions, which are far from the diagonal. However, the diagonal peaks show small phase distortions due to zero-quantum evolution as well and these distortions cause problems for protein NOEs, which are close to diagonal peaks of detergent resonances.

These problems are present in any NOESY spectra. For the bacteriorhodopsin samples, they are however much more pronounced than in spectra of water-soluble proteins, because the artifacts caused by the detergent signals are bigger than the signals of the protein. The main practical problems are the baseline distortions. In indirect dimensions, the problem can be circumvented, if the first data point is recorded with one or the half of the dwell time and back-prediction or first order phase correction are applied (Schmieder, 1991). This is absolutely mandatory for bacteriorhodopsin spectra, especially for the spectral region of the methyl groups.

Otherwise, the artifacts caused by residual protons in the dodecyl chain of the deuterated detergent mask many important spectral details. Additionally it was found, that the amount of artifacts is reduced by oversampling the indirect dimension, e.g. doubling the number points, while recording half the number of FIDs per point and taking half of the dwell time. (Thus, the resolution and the signal-to-noise ratio remain unchanged). Furthermore, anti-diagonals can be placed outside the signal region by setting the receiver frequency to the edge of the spectrum. However, this leads to increased base-line problems in the indirect dimensions.

If the intensities of spectra with different mixing times are to be compared, the long time stability is of crucial importance. A useful approach to circumvent problems in this case is interleaved data acquisition. For each t_1 increment, FIDs of spectra with different mixing times are recorded one after another. Further improvements were achieved by recording a spectrum with zero mixing time and subtracting it from the spectra with longer mixing times, leading to a zero diagonal. For this type of processing, it was found that a separate phase correction of each spectrum before subtraction lead to the best results. Identical processing caused pronounced phase distortions around the diagonal, caused by the zero quantum terms.

During processing, the most crucial step is a careful phase correction, because small phase distortions prevent proper baseline corrections. The main problems are the phase distortions of the diagonal signals of the detergent caused by zero quantum evolution during the mixing time. The phase of these artifacts depends on the difference of the resonance frequencies of the involved spins and is therefore different for different diagonal peaks. It is impossible to find an ideal phase correction of the whole spectrum and the spectra have to be processed with different phase settings to check protein cross-peaks in regions with baseline problems. After the phase correction, automated baseline corrections are performed. The correction routine subtracts a fifth order polynomial from each row or column of the selected region. Baseline corrections were performed over the whole spectral region in both dimensions first. Then additional corrections were performed over smaller regions with the peaks of interest as far as possible from the edge of a region. It was found that correcting regions above and below the diagonal of the spectrum separately lead to the better results as including the diagonal as part of the integration region.

3.4 Proton assignments of bacteriorhodopsin

This paragraph describes assignments of the two forms of dark-adapted bacteriorhodopsin (see chapter 1.4). In what follows the abbreviation 13-cis or cis is used for the protein form with the retinal in 13-cis, 15-syn configuration and all-trans or trans for that with the retinal in all-trans, 15-anti.

3.4.1 Retinal

Starting point for the proton assignments of the all-trans and 13-cis retinals was the H15 resonance of the two isomers Fig. 3.2. The proton resonances of H15-cis and H15-trans have been identified by HMQC spectra on a sample having a $^{13}\text{C}15$ -carbon label. The distinction between 13-cis and all-trans retinal, based on chemical shift analogy to solid state measurements (Smith , 1989b), was confirmed by recording a spectrum on a light adapted sample, which shows the all-trans peak only. The reversibility is shown by a second spectrum run afterwards, in which the two peaks of dark-adapted bacteriorhodopsin reappear (Patzelt , 1997).

NOESY spectra with 5, 10, 20 and 40 ms mixing time have been recorded on a completely deuterated protein sample with protonated retinal. Fig. 3.1 shows parts of the 40ms NOESY spectrum. The assignments of the proton resonances are given in Tab. 3.1, the referencing is relative to the most downfield shifted sugar resonance (5.372 ppm) (Appendix A)

Starting from the H15 resonance of the 13-cis isomer at 8.66 ppm, a line connecting the cross-peaks between neighboring protons of the retinal is drawn in Fig. 3.1. H15 shows a NOE to H14, which shows a strong cross-peak to H20 as well. The other olefinic and methyl protons of the upper part of the retinal are easily assigned as well. The two unresolved methyl groups H16 and H17 are identified via the small cross-peak from H7 to H16/17 shown in the in top part of Fig. 3.1. They are in turn used to identify H8. A remarkable feature of the spectrum is the small cross-peak between H15 and H16/17. The distance between these protons is longer than 7 Å.

The all-trans isomer is assigned in a similar way. The unresolved signals of H12 and H10 are the weakest part of the retinal assignment. The conclusion that H10 and H12 resonances overlap is based on the following arguments:

1. There is only one cross-peak from H14 to the spectral region between 3 and 5 ppm, to a resonance at 3.60 ppm.
2. There is only one cross-peak from H8 to the region between 3 and 5 ppm to the same resonance at 3.60 ppm.
3. H15 shows a small cross-peak to a resonance at 3.58 ppm, which is assigned to H12.

The spectrum in Fig. 3.1 shows the general difficulties of the assignments as well. The region between 3 and 5.5 ppm not shown in the spectrum is dominated by large ridges of t_1 -noise caused by the strong sugar signals of the maltoside. Outside that region, there are a number of NOEs, which are not due to protons of retinal bound to the protein. Some of these resonances have been surrounded by boxes with labels A, B and C.

The three cross-peaks in box A correspond to retinal that is not bound to the protein. The intensity of these signals can be used as indication for sample aging. They become very large in degraded samples. Free retinal can be assigned starting from the H15 resonance at 10.1 ppm, which shows a cross-peak to the methyl group H20 at 2.3 ppm. This methyl shows another cross-peak to H11 at 7.1 ppm. The H19 resonance is at 2.0 ppm and shows cross-peaks to H11 and to H7 at 6.3 ppm. Thus the three cross-peaks in box A are H11-H20, H11-H19 and H7-H19 NOEs of free retinal. The chemical shift of the protons in micelles are very similar to reported ^1H chemical shifts for all-trans retinal in hexamethyldisiloxane at 32°C (Rowan III, 1972; Rowan III and Sykes, 1975) or in CDCl_3 at 25°C (Pattaroni, 1981) and at 44°C (Patel, 1969).

The resonances in the box B are resulting from spurious amounts of Triton X-100, a detergent used during sample preparation (Patzelt, 1997). These peaks are in a region of the spectrum that contains NOEs between aromatic protons and methyl groups in samples with different labelling patterns. Therefore the Triton peaks are masking important information in some of the spectra. The presence of the Triton peaks can be avoided by solubilizing the purple membrane directly in dodecyl-maltoside, a method applied in the preparation of some of the samples.

The resonances in box C are in a spectral region, which suffers from artifacts of the strong sugar signals of the detergent. The row of peaks from 6.0 ppm (F2) to 4.6, 3.7 and 3.4 ppm (F2), might be due to cross-peaks between the β -anomeric proton and other sugar resonances of a detergent molecule that is tightly bound to the protein.

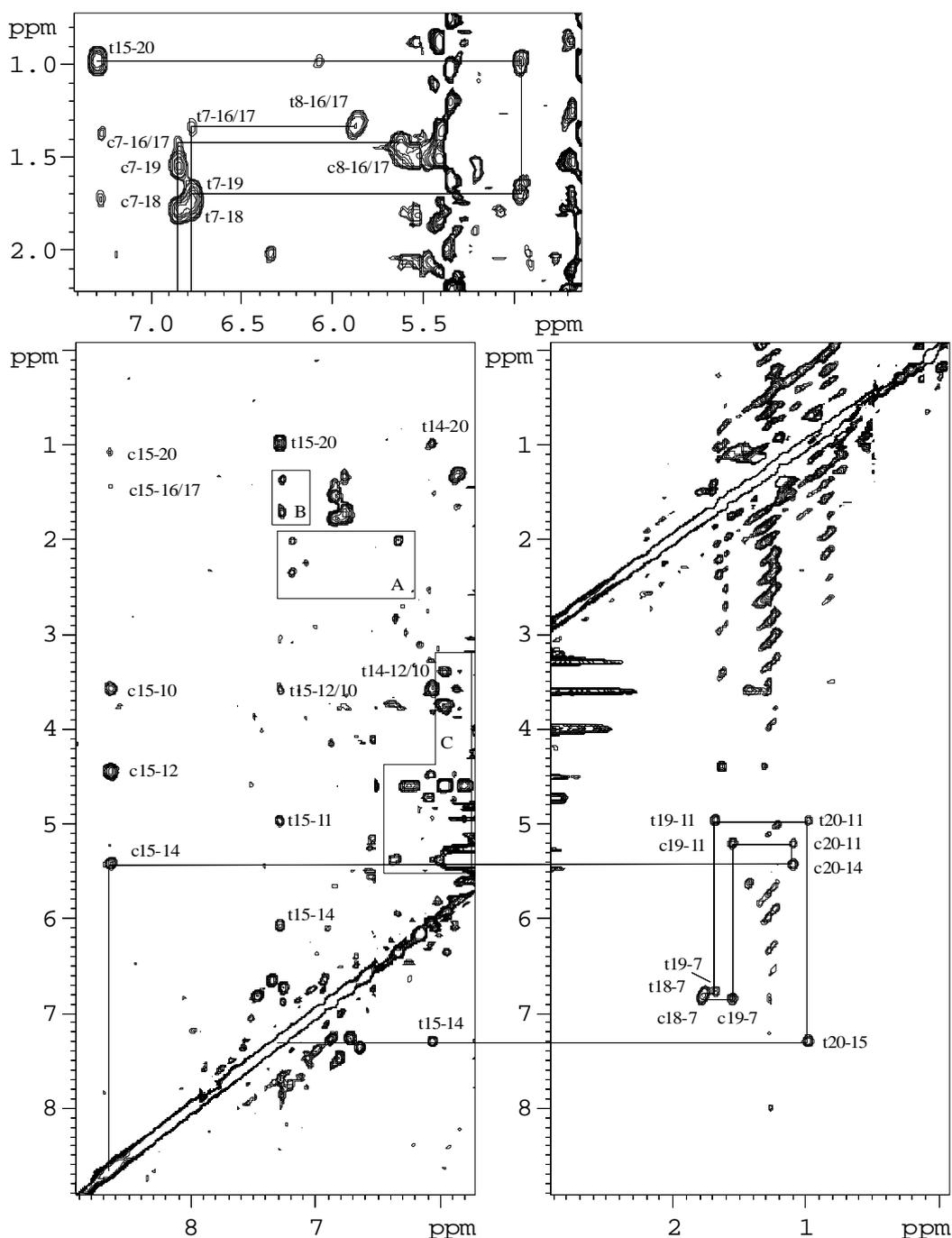


Fig. 3.1 NOESY spectrum of ^1H retinal, ^2H bacteriorhodopsin with a mixing time of 40ms. A small t / c in front of an assignment indicates that the resonance is for the all-trans / 13-cis isomer of retinal, e.g. c15-12 indicates a cross-peak between H15 (horizontal axis) and H12 (vertical axis). The lines connecting some peaks are there to guide the eye through part of the sequential assignments. The inserted boxes indicate peaks of other molecules present in the sample (see text).

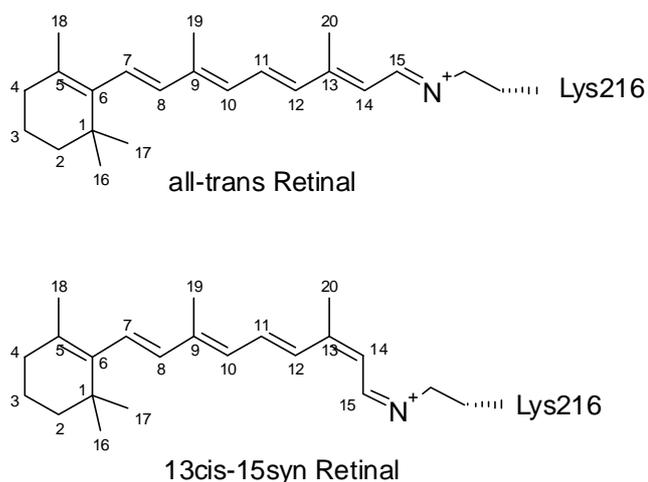


Fig. 3.2 Numbering of the carbon atoms in the two retinal isomers in dark-adapted bacteriorhodopsin.

	H7	H8	H10	H11	H12	H14	H15
13-cis	6,84	5,62	3,58	5,21	4,45	5,43	8,64
all-trans	6,77	5,87	3,58	4,98	3,60	6,08	7,28
	H16	H17	H18	H19	H20		
13-cis	1,44	1,44	1,78	1,54	1,10		
all-trans	1,35	1,35	1,75	1,68	0,98		

Tab. 3.1 Chemical shift of the retinal protons of the all-trans and 13-cis isomers in dark-adapted bacteriorhodopsin.

3.4.2 Threonine and tryptophan assignments

The specific assignments of selected amino acid proton resonances are much less straight forward than the retinal assignments. There is only one retinal - in two different isomeric forms - per protein, but there are 8 Trp, 18 Thr, 11 Tyr, 9 Met and 7 Lys. The chemical shift dispersion of equivalent protons in these amino acids is not expected to be very big in general and the large linewidth of the signals (Appendix D) leads to substantial overlap.

Four of the 8 Trp residues are part of the retinal binding pocket (Trp86, Trp138, Trp182 and Trp189). If spectra of ^1H Trp labelled samples are processed with a very strong resolution enhancement, the signals may be clustered to eight AMXP spin-systems. For some of these clusters, it is possible to distinguish between the outer ($\text{H}\zeta_2$, $\text{H}\epsilon_3$) and inner ($\text{H}\zeta_3$, $\text{H}\eta_2$) protons of the indole ring (see Fig. 3.3) based on observed cross-peak intensities in spectra of short mixing times. However, the identification of the clusters is far from unique because of large overlap of most of the aromatic resonances. Two of the aromatic protons (5.97 and 6.15 ppm) resonances are

shifted about 1 ppm upfield compared to the bulk of resonances, which are between 6.7 and 7.7 ppm. 12 AMX spin systems for Thr can be identified by cross-peaks from the methyl-group (acquisition dimension) and the α and β protons (indirect dimension). Some of the identified clusters might be due to more than one Thr. The intensities in the spectral region of the cross-peaks are strongly disturbed by t1 noise and it is therefore not always possible to assign H α and H β uniquely.

For the sequence specific assignments some information from structural models was used. The initial atomic model by Henderson et al. (Henderson, 1990) is sufficient for the assignments reported in this paragraph. The electron crystallographic resolution of the data used to generate the model was 3.5 Å in the plane of the membrane and approximately 10 Å in the direction perpendicular to it. Data with such a resolution allow to resolve the side-chains of Trp unambiguously and therefore contain the additional information needed for NMR assignments at this stage. The higher resolution structures published later (Essen, 1998; Grigorieff, 1996; Luecke, 1998; Pebay-Peyroula, 1997) helped mainly to resolve the problem in assigning the threonine side-chain closest retinal protons H20 and H11. In this case the two neighboring Thr 89 and 90 were possible candidates. However, swapping assignments for Thr 90 to Thr 89 could be ruled out by comparing the calculated structures with alternate assignments. Most assignments have been made before the first x-ray structure (Pebay-Peyroula, 1997) was published. All assignments in this paragraph include a network of NOEs with retinal protons, for which specific assignments were obtained without information from crystal structures.

The sequence specific assignments were started based on Ret to Trp NOEs. In the aromatic region a NOE from H15 CIS (8.64 ppm) to a resonance at 6.64 ppm is observed. This peak shows up in spectra of 10 ms mixing time. Based on the crystal structures, the resonance at 6.64 ppm was identified as a proton of Trp86. In spectra with longer mixing times no additional NOE between H15 CIS and other aromatic protons is observed, which suggests the identification of the proton as H δ 1. If the proton at 6.63 ppm would belong to an AMXP spin system additional peaks at longer mixing times would be expected due to spin-diffusion. The assignment is confirmed by additional NOEs to the retinal protons H12 and H14 and to the methyl protons of Thr 90.

In spectra with a mixing time of 80ms, cross-peaks between the Retinal methyl groups H16/H17 and H20 to several protons, which are part of an AMXP spin systems, are observed. Based on the crystal structures, it was assumed that H16/H17 are closest to Trp86 and H20 is closest to Trp182. The spin-systems close to H16/H17 CIS and TRANS include the two strongly upfield shifted aromatic proton resonances at 5.96 ppm and at 6.13 ppm. For these two quite well resolved aromatic spin-systems, it is possible to differentiate between the outer protons ζ_2 and ϵ_3 and the inner protons ζ_3 and η_2 of the AMXP spin systems based on the patterns and build-up curves of the intra-residue cross-peaks. This information together with the inter-residue NOEs between these resonances and the retinal allows a complete assignment of all aromatic protons of Trp86 Cis. For Trp86 Trans, the resonances show more overlap and the situation is less clear. The all-trans Trp86 H δ_1 proton is assigned by its cross-peak to H14 and the other protons on the assumption that identical protons in Cis and Trans have similar chemical shifts. The specific assignments for all-trans are confirmed by semi-empirical proton chemical shift calculations and by calculation of NOE build-up curves from the coordinates (see chapter 3.5 and 3.6)

There is one cross-peak from H16/17 Cis and Trans to a Trp proton, which does not belong to the spin-system of Trp86. Based on information from the crystal structures and chemical shift calculations, this proton is assigned to Trp189 H δ_1 .

All-trans and 13-cis Trp 182 are identified by NOEs to the retinal methyl H20 of both isomers. It shows cross-peaks to the two most upfield shifted threonine methyl groups and probably one Met methyl group. One of the two Thr methyl groups shows in turn NOEs to Ret CH₃20, CH₃19, H11, H15 (all-trans and 13-cis), H14 and H10 (only 13-cis). The chemical shift of the corresponding all-trans and 13-cis threonine residues are very similar for all protons. The cross-peak patterns and the chemical shift allow the unique identification of Thr 90 and Thr 178. The specific assignments of the four protons of the Trp 182 indole rings are based on the NOE patterns and the distances observed in the high-resolution X-ray structures. These assignments cannot be made on the base of the NMR data alone, since the resonances are overlapped. In the structure calculations, the four protons were treated as not specifically assigned. For distance constraints, a pseudo-atom located in the geometric center of the four protons is introduced.

One more set of NOEs between the three resonances of a threonine residue and Ret CH₃18 is observed in all spectra of samples containing ¹H Thr and ¹H Ret. The only threonine close to Ret CH₃18 is Thr142. This assignment is confirmed by the observation of four NOEs between the methyl of Thr142 and aromatic tryptophan protons, which are in turn assigned to be the AMPX spin system of Trp138. Specific assignments of the four protons of the indole ring are not possible considering the NOESY spectra. The assignments are based on cross-peaks observed in NOESY spectra with mixing times equal or longer than 40ms.

All assignments are summarized in Tab. 3.2

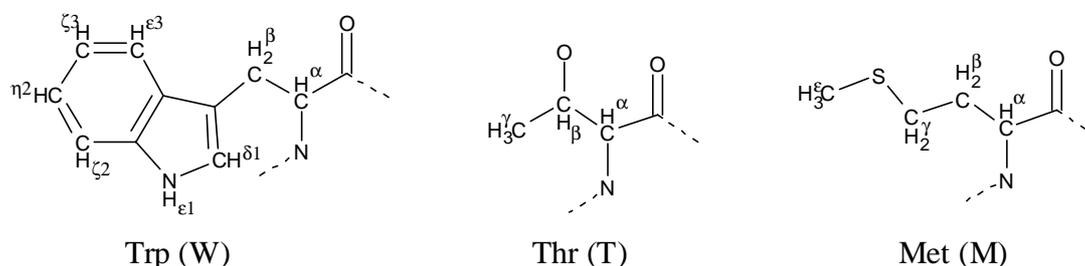


Fig. 3.3 Naming convention used for tryptophan, threonine and methionine protons.

	H α	H β	H γ		
Thr90 Cis	4.24	4.56	0.28		
Thr90 Trans	4.26	4.60	0.22		
Thr142	3.63	3.43	0.83		
Thr178 Cis	3.87	4.08	-0.05		
Thr178 Trans	3.77	4.00	-0.13		
	H δ 1	H ϵ 3	H ζ 3	H ζ 2	H η 2
Trp86 Cis	6.64	6.99	6.64	7.06	5.97
Trp86 Trans	6.88	7.03	6.73	6.91	6.15
Trp138			6.93*	7.11*	7.00*
Trp182 Cis	7.43	7.52*	7.07*	7.45*	7.10*
Trp182 Trans	7.38	7.38*	7.05*	7.27*	7.19*
Trp189 Cis	7.50				
Trp189 Trans	7.41				

Tab. 3.2 Chemical shift assignments of the proton resonances of amino-acid residues used in the structure calculation. The * indicates that the corresponding assignments might have to be interchanged.

3.4.3 Assignments based on published high resolution structures

The assignments given in the current paragraph rely much more on existing high-resolution structures. This is because the NOE networks do not contain any retinal

resonances and the number of possibilities is therefore much larger. Additional cross-peaks identified as Thr-Trp or Met-Trp NOEs are observed in the spectra. The possible assignments are based on considering all Met and Thr residues, which are close enough in space to the aromatic protons of a Trp side-chain. The radius around a proton was chosen generously (e. g. up to 8 Å) to include possible distance errors for pairs of residues in some of the structures considered. For all Thr-Trp NOEs, a consistent picture is obtained in this way, meaning that all NOEs could be assigned uniquely (including the information in chemical shift calculations of the next chapter).

Two strong cross-peaks between a Thr CH₃ at 1.52 ppm and aromatic protons at 7.50 and 6.85 ppm are present in spectra with short mixing times. At longer mixing times, peaks at 6.99 and 7.27 ppm are observed to the same methyl group. The H α (4.60 ppm) and H β (3.97 ppm) of the same Thr show NOEs to the same aromatic protons. Distances determined from the volumes of the three peaks at 7.50 ppm in a NOESY spectrum with 10 ms mixing time are 3.06 Å (to Thr H γ), 2.21 Å (to Thr H β) and 2.72 Å (to Thr H α). The cross-peak at 6.85/1.52 ppm is overlapping with intra retinal cross-peaks and the two cross-peaks 6.85/3.97 and 6.85/4.60 are observed in spectra with longer mixing times only. In the crystal structures, only four threonine residues are closer than 5 Å to more than one aromatic tryptophan proton. Three of these have been identified as Thr 90, Thr 142 and Thr 178. The NOEs can be assigned with high confidence to the fourth pair, namely Thr 121 and Trp 137.

Another set of strong NOEs between three Thr protons and one aromatic Trp proton is observed in spectra with a mixing time of 5 and 10 ms. The derived distances are 2.97 Å (H γ , 1.26 ppm), 2.41 Å (H β , 4.19 ppm) and 3.89 Å (H α , 4.81). The α proton at 4.81 ppm is the most downfield shifted threonine resonance observed in the spectra. Its chemical shift falls in a region, where H α protons of threonines in β -sheet secondary structures are expected. The loop between helices E and F forms a short two stranded β -sheet structure (Kimura, 1997; Luecke, 1999b). Thr 67 is part of this β -sheet and the distances to Trp 80 correspond well to the observed cross-peak pattern (e.g. W80H δ 1-T67H γ 2.84 Å -T67H γ 2.19 Å -T67H γ 4.02 Å in 1c3w (Luecke, 1999b)).

Two to three different protons of methionine residues show NOEs to aromatic tryptophan protons. They probably are originating from the methionine methyl group, which has more favorable relaxation properties than the methylene groups (Appendix D). Moreover, if a methionine methylene group would show such strong NOEs in spectra with short mixing times, more cross-peaks to the neighboring methionine protons would be expected in spectra with longer mixing times. The chemical shifts of the Met/Trp NOEs are 1.98 ppm to 7.52, 7.06 and 6.81 ppm and 2.11 ppm to 7.42, 6.76, 7.27 and 7.00 ppm. A third row of resonances is in a more crowded region of the spectrum at 1.66 ppm, at the same chemical shift than CH₃19 Trans and only the changed intensity of some NOEs compared to a 1H Trp/1H Ret labelled sample indicates possible additional Trp/Met cross-peaks. In the crystal structures, the only distances of Met methyl to Trp aromatic protons short enough to expect observation of NOEs are between the pairs W12/M209, W80/M56 and W182/M118 and M145. The only distances shorter than 3 Å in the coordinate files is Met209Hε-Trp12Hε3 (Belrhali, 1999; Luecke, 1999b).

Additional information about methionine chemical shifts can be obtained from the literature. Seigneuret et al. (Seigneuret, 1991a) recorded ¹³C detected ¹³C-¹H COSY spectra of bacteriorhodopsin with methyl-¹³C labelled methionine solubilized in Triton X-100 at pH 5 and 20° C. They identified 4 resolved ¹³C peaks at 17.2, 15.4, 14.9 and 14.0 ppm with intensity ratios 3:4:1:1. The bound protons resonate at 2.2 ppm (¹³C shifts 15.4 and 14.9 ppm) and at 1.7 ppm (¹³C shift at 17.2ppm); (the method of referencing the proton chemical shift in the spectrum is not mentioned.) The ¹³C resonance at 14.0 ppm shows no cross-peak in their ¹H-¹³C correlation spectrum. The authors provide some hints for the assignment of the resonances based on observing chemical shift changes upon addition of hydrophilic and amphiphilic paramagnetic reagents and based on papain proteolysis. Their conclusions are: (1) The cross-peak 15.4 ppm ¹³C/ 2.2 ppm ¹H contains four resonances, one of which is Met68. (2) The hydrophilic paramagnetic reagent affects the resonances 15.4 ppm ¹³C/ 2.2 ppm ¹H and 14.9 ppm ¹³C/ 2.2 ppm ¹H. (3) The amphiphilic paramagnetic reagent affects the same two resonances and additionally the 17.2 ppm ¹³C / 1.7 ppm ¹H signal.

The random coil shift of the Met CH₃ is 2.13 ppm (Wüthrich, 1986). The calculated chemical shifts are 1.77 ppm (Met 56 H ϵ), 1.88 ppm (Met 118 H ϵ), 1.08 ppm (Met 145 H ϵ) and 2.39 ppm (Met 209 H ϵ) (coordinates 1c3w (Luecke, 1999b), see chapter 3.5 for details concerning chemical shift calculations).

Taking all observations together, the following interpretation of the Trp/Met NOEs is favored: The four peaks at 2.11 ppm are from CH₃ ϵ Met209 to Trp12. The row of NOEs at 1.98 ppm is from Met118 to Trp182 Cis. The corresponding Met118 Trans resonance is at 1.7 ppm and thus at almost the same frequency as Ret CH₃19 Trans. This interpretation would be confirmed, if the row of peaks at 1.98 ppm disappears in a spectrum of light adapted bacteriorhodopsin.

3.5 Analysis of the chemical shifts

3.5.1 Introduction of the theoretical model

Computer programs for the calculation of proton chemical shifts from a protein structure have been developed over the last years (see chapter 1.3). In the present paragraph, we use the program SHIFTS (Version 3.0b2) (Ösapay and Case, 1991) to calculate chemical shifts from the structures of the all-trans ground state of bacteriorhodopsin deposited in the Brookhaven protein data base (pdb). The program calculates the chemical shifts of protons as the sum of the following contributions:

1. Ring-current contributions are calculated according to Haigh-Mallion theory (Haigh and Mallion, 1980).
2. Magnetic anisotropy of the peptide groups are calculated according to the model of McConnell (McConnell, 1957) under the assumption of axial symmetry. The origin of the anisotropy is chosen 0.7 Å from the carbon along the bisector of the NCO angle. Since the peptide anisotropy contributes to the random coil shift as well as to the proton shift two additional constants are added to correct the average C α proton and the average side-chain proton shifts.
3. Electrostatic effects are calculated to be proportional to the projection of the local electric field on the C-H bond vector. The local electric field is determined by considering the backbone charges only, which are taken from the charge model of the CHARMM (version 19) parameters (-0.35, 0.25, 0.20, 0.55 and -0.55 for N, H, Ca, C, and O atoms, respectively, and -0.20 for the charge of a proline nitrogen).

The final model in SHIFTS has 11 parameters: seven ring-current intensity factors (two of which are considered in heme proteins only, which were included in the

analysis in (Ösapay and Case, 1991)), one parameter each for the peptide group anisotropy and for the backbone electrostatic contributions, and two constants, one for C α protons and one for side-chain protons. These parameters have been adjusted by minimizing the difference between calculated and observed chemical shifts of 20 proteins, for which crystal structures and NMR assignments were available. A total chemical shift rms of 0.220 for all protons in the observed non-heme proteins was achieved. Shifts of side chain protons (rms=0.203) and the methyl groups (rms=0.136) show a higher precision, the amides shifts show the largest deviations (rms=0.575 for all amides and r=0.506 for those involved in internal hydrogen bonds).

The ring-current intensity factors and the factor scaling the electrostatic contributions have been calibrated later by quantum chemical shielding calculations for methane molecules placed in a variety of positions near the aromatic rings of the type found in proteins (Case, 1995). In general, the solutions found by the density functional calculation lead to larger parameters than the ones obtained from the empirical fits.

For all calculations used in what follows, the older empirical parameters of (Ösapay and Case, 1991) are used. This is in accord with the current release of the program SHIFTS (V.3.0b2). The larger quantum chemical calibration factors resulted in differences of up to 1 ppm for the most strongly shifted signals in bacteriorhodopsin. The experimental results are reproduced much better by using the old calibration factors.

Out of the eleven optimized parameters of SHIFTS, the following eight are used for the calculation of bacteriorhodopsin chemical shifts are (the three parameters not needed are the ring-current intensity factors for heme and histidine):

1. Four ring-current intensity factors: 1.00 for Phe, 0.84 for Tyr, 1.04 for Trp-5 and for 1.02 Trp-6
2. Peptide groups anisotropy: $-7.9 \text{ erg/G}^2\text{-mol}$
3. Proportionality factor for backbone electrostatic contributions: $-2.69 \pm 0.19 \cdot 10^{-12} \text{ esu}$
4. Constant for C α protons: -0.754
5. Constant for side-chain protons: -0.041

To compare the calculated secondary shifts to the experimentally determined values, the following random coil values are added in SHIFTS (V. 3.0b2):

1. Thr: H α 4.39, H β 4.32, CH₃ γ 1.22
2. Trp: H δ 1 7.22, H ζ 2 7.44, H η 2 7.19, H ζ 3 7.13, H ϵ 3 7.59

Using the slightly different values for the random coil shifts given by Wishart et al. (Wishart, 1995) lead to similar results as those presented below.

The Retinal atom coordinates were not considered in the calculation of amino acid chemical shifts. This is consistent with the model used for amino acid side chains, the charge of which is not used during the calculation. The sequence of the protein entering the calculation specifies residue 216 as a lysine. The program does model the missing protons to the pdb-input file and ignores any atoms it doesn't recognize from the sequence.

3.5.2 Comparison of the experimental chemical shifts with shift calculations based on a X-ray structure at 1.55 Å resolution

In Tab. 3.3 the shifts calculated from x-ray coordinates at 1.55 Å resolution (1c3w, (Luecke, 1999b)) are compared to the experimental values of all-trans bacteriorhodopsin. The overall rms between calculated and observed values is 0.191 for all assigned protons and 0.206 for the protons surrounding the retinal. The second subset (W86, W182, T90 and T142) is chosen for additional comparison, because we assume that the conformation of the side-chains of these residues is not influenced significantly by crystal contacts in the x-ray structure and the experimental chemical shifts are not influenced by protein detergent interactions. In general, the correspondence between calculated and observed chemical shifts is very high for the Thr methyl groups and the aromatic Trp residues. H α and especially the H β protons of Thr show larger deviations.

amino acid	number	atom	exp. Shift	cal. Shift	Exp.-cal.
Thr	67	H α	4.81	4.74	0.07
Thr	67	H β	4.19	4.43	-0.24
Thr	67	H γ	1.26	1.23	0.03
Thr	90	H α	4.26	4.09	0.17
Thr	90	H β	4.60	3.99	0.61
Thr	90	H γ	0.22	0.36	-0.14
Thr	121	H α	3.97	4.01	-0.04
Thr	121	H β	4.60	4.69	-0.09
Thr	121	H γ	1.52	1.41	0.11
Thr	142	H α	3.63	3.68	-0.05
Thr	142	H β	3.43	3.81	-0.38
Thr	142	H γ	0.83	0.76	0.07
Thr	178	H α	3.77	3.93	-0.16
Thr	178	H β	4.00	4.26	-0.26
Thr	178	H γ	-0.13	-0.37	0.24
Trp	86	H δ 1	6.88	6.82	0.06
Trp	86	H ζ 2	6.91	6.85	0.06
Trp	86	H η 2	6.15	6.26	-0.11
Trp	86	H ζ 3	6.73	6.85	-0.12
Trp	86	H ϵ 3	7.03	7.18	-0.15
Trp	182	H δ 1	7.38	7.23	0.15
Trp	182	H ζ 2	7.27	7.21	0.06
Trp	182	H η 2	7.19	7.11	0.08
Trp	182	H ζ 3	7.05	6.92	0.13
Trp	182	H ϵ 3	7.38	7.37	0.01
rms (all protons)			0.191	r (all)	0.820
rms (core: T90, 142, W86, 182)			0.206	r (core)	0.767
rms (aromatic/methyl)			0.118	r (arom/met)	0.933

Tab. 3.3 Comparison of calculated and experimental chemical shifts for all-trans bacteriorhodopsin. The calculated shifts are derived from a x-ray structure with 1.55 Å resolution (accession number 1C3W in the Protein Data Bank) (Luecke, 1999b) using the program SHIFTS v3.0b (Case, 1995; Ösapay and Case, 1991). The rms (root mean square deviation) and the linear correlation coefficient are defined in Eq.(3.1) and Eq.(3.2), with x_i and y_i being the deviation of the observed/calculated shifts from their random coil values.

The overall rms between observed and calculated shifts is defined as

$$\text{rms} = \sqrt{\frac{1}{N} \sum (x_i - y_i)^2} \quad (3.1)$$

The relation between calculated and observed secondary shift can be analyzed by evaluating the experimental linear correlation coefficient r , which is defined as

$$r \equiv \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{(N \sum x_i^2 - (\sum x_i)^2)(N \sum y_i^2 - (\sum y_i)^2)}} \quad (3.2)$$

where N is the number of observation points x_i ($\Delta\delta_i^{\text{obs}}$) and y_i ($\Delta\delta_i^{\text{cal}}$). Values of r range from 0, when there is no correlation, to ± 1 for complete correlation. More quantitative statements for the significance of r are obtained under the assumption of a

Gaussian distribution. The probability distribution of r is compared to a completely uncorrelated parent distribution. Integration of the so obtained distribution P from $|r|$ to 1 gives the probability that a random sample of N uncorrelated data points would yield an experimental linear-correlation coefficient as large as or larger than the observed value for $|r|$. This probability is (Bevington and Robinson, 1992)

$$P_c(r; N) = \frac{2}{\pi} \frac{\Gamma((v+1)/2)}{\Gamma(v)} \int_{|r|}^1 (1-r^2)^{(v-2)/2} dr \quad v = N - 2 \quad (3.3)$$

Equation (3.3) cannot be integrated analytically. Numerical integration was performed using the program Mathematica (Wolfram, 1996). Values for P_c of 2.9×10^{-7} ($r=0.820$, $N=28$), 5.3×10^{-4} ($r=0.767$, $N=16$) and 1.34×10^{-7} ($r=0.933$, $N=16$) are obtained. The probability that the null hypothesis is true is less than 0.1 % in all cases, indicating that it is extremely improbable that the calculated and measured secondary chemical shifts are linearly uncorrelated. It should be pointed out that equation (3.3) is based on the assumption that the distributions for x and y form a two-dimensional Gaussian distribution around their mean values. If this assumptions doesn't hold, the probability interpretation of the null hypothesis are meaningless (Press, 1988).

The experimental correlation $r=0.820$ and $rms=0.191$ for all assigned protons are slightly better than values obtained by Ösapay and Case ($r=0.880$, $rms=0.233$) for a much larger statistical sample of 5678 protons (Ösapay and Case, 1991). The largest differences found are for the threonine H β shifts. This might be due to the influence of the hydrogen-bonding contribution of the neighboring side chain oxygen. In simple molecules and in the absence of other effects, hydrogen bond formation leads to downfield shifts (higher ppm values) of the resonances of both amide protons and carbonyl carbons (Jardetzky and Roberts, 1981). The carbon deshielding indicates a reduced electron density, caused by the polarization of the CO bond. In proteins, amide shifts are influenced by other factors as well. However, it still seems that strong internal hydrogen bonds lead to downfield and long weak hydrogen bonds to upfield shifts compared to the random coil values in water (Wishart and Sykes, 1994). For the threonine side chain, it would be expected that a strong hydrogen bond of the side-chain oxygen would repel electrons from the β carbon and cause a downfield shift of both the carbon and β proton resonances. A closer look at the deviations of the

observed chemical shifts of Tab. 3.3 is not in contradiction to this qualitative picture. With the exception of one (Thr 90) out of five threonines, the calculated shifts are on average 0.24 ppm larger than the observed ones. Interpreting this shift being caused by differences of hydrogen bond strength only indicates that the side-chain oxygen of T67, T121, T142 and T178 are less hydrogen bonded in the protein than in aqueous solution.

It remains to explain the large deviation between calculated and observed chemical shift of Thr 90 H β , which is not only large in magnitude but also of opposite sign than all other observed H β shift deviations. The side-chain of Thr 90 should show a very strong hydrogen bond, if other effects contributing to the chemical shift could be neglected. In the 1.55 Å X-ray structure (Luecke, 1999b) Thr90 (OG1) is indeed showing a hydrogen bond to Asp115 (OD1). However, due to the close proximity of Thr90 to the conjugated π system of the retinal, electric charge and magnetic susceptibility effects are of importance, although to small to account for the observed large difference.

We may conclude, that the effect of hydrogen bonding could explain the observed large shift deviations for the H β protons of Thr based on a simple qualitative argumentation. The observation of just 5 shifts is of course by no means significant and the hydrogen bonding effect should be confirmed, e. g. by a systematic search of a data base with protein assignments (BioMagResBank) (Peavey, 1991). A confirmation would make the T90 H β shift to a direct and easy accessible indicator for the hydrogen bonding state of its side-chain, which could be very useful for measurements of intermediate structures in the photo-cycle.

The distribution of all identified AMX spin-systems and calculated secondary chemical shifts for the threonine CH $_3\gamma$ are shown in Fig. 3.4. The figure illustrates that sequence specific assignments are possible from a comparison of observed and calculated shifts alone. The two methyl groups showing the largest upfield shifts are assigned to Thr178 and Thr90. The third largest observed upfield shift is not as clear, but if the distribution of the Thr H α protons is considered as well, Thr142 can be uniquely identified from the calculated chemical shifts.

A similar analysis of the tryptophan chemical shifts would allow to identify the AMXP spin-system of Trp 86. Its H η 2 is the most upfield shifted aromatic resonance in the calculation.

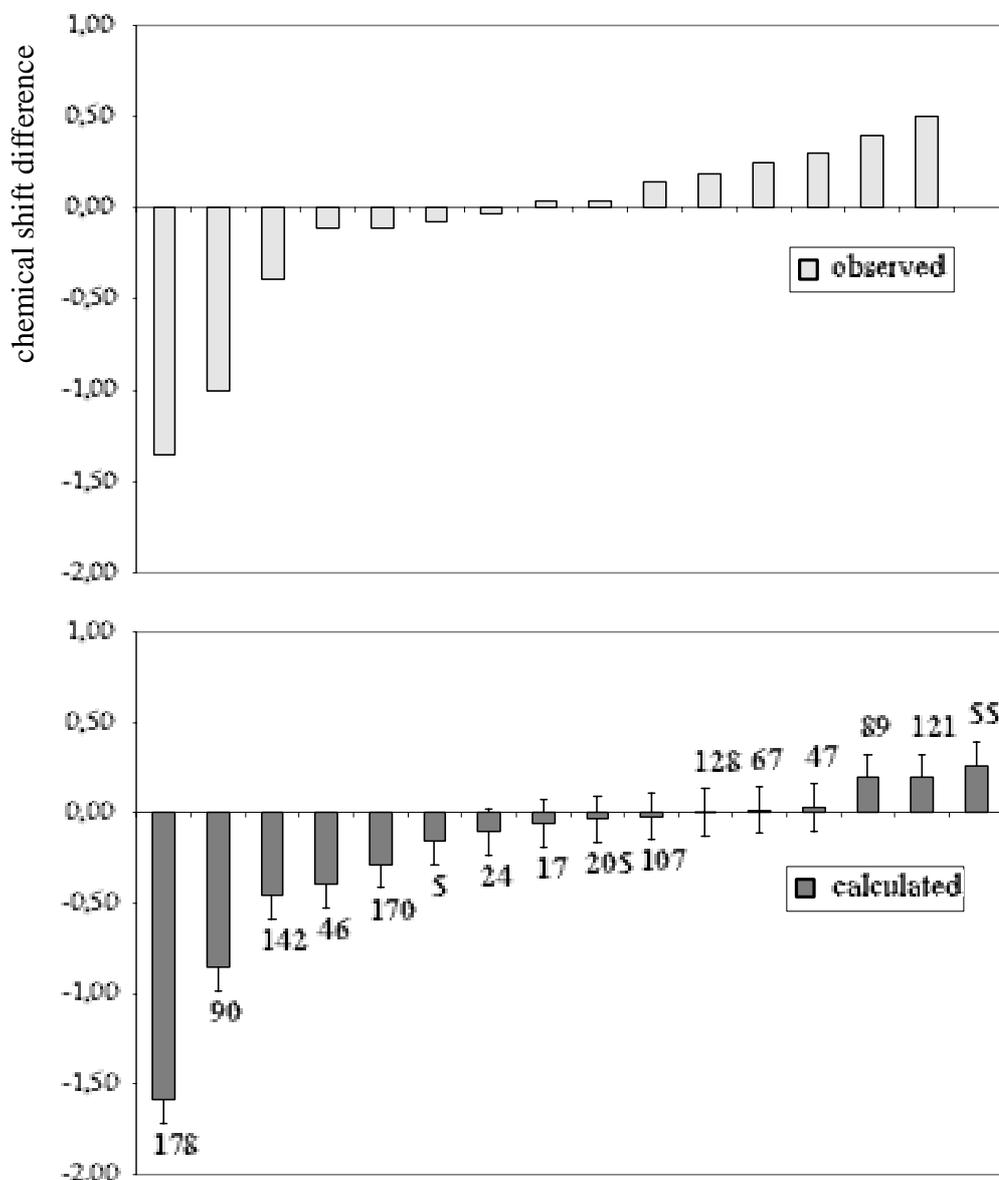


Fig. 3.4 Observed (top) and calculated (bottom) chemical shift differences from the random coil value for all threonine residues in bacteriorhodopsin. The calculations use the coordinates of a X-ray structure at 1.55 Å resolution (pdb-entry 1c3w (Luecke, 1999b)), the error bars correspond to the rms for all methyl groups of the proteins in the calibration data set of the program used for calculating the shifts (Ösapay and Case, 1991). The observed chemical shifts of all AMX spin systems identified from the threonine spectra. Three of the five sequence specific assignments obtained from analyzing networks of NOEs could be obtained directly from this figure (see text).

It can be concluded, that it is possible to assign Thr178, Thr142, Thr90 and the AMXP spin-system of Trp86 by comparing observed and calculated chemical shifts

alone. The theoretical model is very precise in calculating these shifts and the observed shifts are separated from other resonances of the same amino acids, because the structural shifts of the H γ protons of these residues are the largest of all threonine residues.

3.5.3 Chemical shifts as a tool to analyze the accuracy of a structure

Over the last three years, the coordinates of a number of high-resolution structures of the all-trans ground state have been published and are available in the Brookhaven protein database. Seven different research groups using two different methods (electron or X-ray crystallography) contributed data. The first structure with a resolution of 3.5 Å in the plane of the membrane and of 4.3 Å in the perpendicular direction ((Grigorieff, 1996), pdb entry code 2brd) was obtained by electron crystallographic refinement on an earlier model, which was directly based on the experimentally observed density (Henderson, 1990). A *de novo* structure determination based on image phases ((Kimura, 1997), pdb entry code 1at9) and later refined using electron crystallographic data ((Mitsuoka, 1999), pdb entry code 2at9) achieved a resolution of 3.0 Å. All X-ray structures published to date are refined by molecular replacement, with the electron microscopic structures as starting models. The first X-ray structure (Pebay-Peyroula, 1997) was obtained with a novel approach for obtaining three-dimensional crystals using lipid cubic phases (Landau and Rosenbusch, 1996). The same type of crystals were analyzed by Luecke et al. ((Luecke, 1998), pdb entry code 1brx), who realized the presence of substantial hemihedral twinning. This lead Pebay-Peyroula and coworkers to repeat their refinement subsequently (pdb entry code 1ap9) (see discussion in (Subramaniam, 1999)). Structures of different resolution were published subsequently, either to discuss more structural details of the all-trans ground state ((Belrhali, 1999), pdb entry code 1qhj) ((Luecke, 1999b), pdb entry code 1c3w) or as data for comparison with intermediates of the photo-cycle ((Luecke, 1999a), pdb entry code 1c8r: all-trans and 1c8s: M intermediate, (Sass, 20-OCT-99), pdb entry code 1cwq molecule A: all-trans, molecule B: M intermediate, (Edman, 1999), pdb entry code 1qkp: mixture of ground state and K intermediate and 1qko: refined K intermediate). In addition to these structure obtained from the same crystal form, two different types of crystals have

yielded high-resolution structures ((Essen, 1998), pdb entry code 1brr (Sato, 1999), pdb entry code 1bm1). In the monoclinic bacteriorhodopsin crystals of Essen et al., there are three molecules per unit cell, leading to three slightly different structures, which will be referred to as 1brr_A, 1brr_B and 1brr_C. A detailed comparison of the different structures published until the end of 1998 and a discussion of their implication on the function of the protein are found in the review of Subramaniam (Subramaniam, 1999).

In Tab. 3.4, the chemical shifts calculated from the highest resolution structures for each of the research groups, which published ground state structures, are listed. The average over the six values and their standard deviations are included as well. It can be seen that the high similarity of the published structures is reflected in the calculated chemical shifts, which differ by less than 0.2 ppm for most of the protons. Larger standard deviations of the average shifts are observed as soon as one of the crystallographic structures deviates significantly from the others. The most striking example is T178. In the structure of Essen et al., the side-chain of this threonine is clearly in a conformation, which does not agree with the observed chemical shifts. The calculated structural shift of the methyl group differs by more than 1ppm from the observed and the calculated shifts of all other structures. For crystallographic studies the differences in the electron density of the methyl and hydroxyl groups attached to the β carbon of a threonine side chain is very small. The decision which rotameric state is chosen for the side-chain dihedral angle χ_1 is often based on considering possible electrostatic interactions in all orientations, which fit the electron density equally well. Choosing a similar χ_1 in 1brr as observed in other X-ray structures fits the electron density equally well (L. O. Essen personal communication).

proton	exp.	1c3w	1qhj	1br	1bm1	2at9	2brd	av.	std.
T67 H α	0.42	0.35	0.15	0.10	0.44	0.01	0.22	0.21	0.16
T67 H β	-0.13	0.11	0.09	-0.05	0.11	0.16	0.11	0.09	0.07
T67 H γ	0.04	0.01	0.05	-0.12	-0.08	0.56	-0.11	0.05	0.26
W86 H δ 1	-0.34	-0.40	-0.37	-0.30	-0.43	-0.45	-0.37	-0.39	0.05
W86 H ζ 2	-0.53	-0.59	-0.51	-0.63	-0.85	-0.43	-0.60	-0.60	0.14
W86 H η 2	-1.04	-0.93	-0.97	-1.07	-0.96	-0.64	-0.69	-0.88	0.17
W86 H ζ 3	-0.4	-0.28	-0.39	-0.59	-0.67	-0.39	-0.07	-0.40	0.22
W86 H ϵ 3	-0.56	-0.41	-0.44	-0.61	-0.71	-0.47	-0.23	-0.48	0.17
T90 H α	-0.13	-0.30	-0.26	-0.31	-0.48	-0.40	-0.52	-0.38	0.11
T90 H β	0.28	-0.33	-0.31	-0.31	0.04	-0.17	0.07	-0.17	0.18
T90 H γ	-1	-0.86	-0.87	-0.86	-0.56	-0.58	-0.46	-0.70	0.19
T121 H α	-0.42	-0.38	-0.45	-0.71	-0.26	-1.11	-0.71	-0.60	0.31
T121 H β	0.28	0.37	0.31	0.37	0.46	0.14	0.26	0.32	0.11
T121 H γ	0.3	0.19	0.16	0.10	0.12	0.15	0.02	0.12	0.06
T142 H α	-0.76	-0.71	-0.62	-0.65	-0.76	-0.57	-0.70	-0.67	0.07
T142 H β	-0.89	-0.51	-0.39	-0.37	-0.29	-0.26	0.03	-0.30	0.18
T142 H γ	-0.39	-0.46	-0.43	-0.40	-0.41	-0.38	-0.40	-0.41	0.03
T178 H α	-0.62	-0.46	-0.35	-0.27	-0.67	-0.11	-0.99	-0.48	0.31
T178 H β	-0.32	-0.06	-0.05	-1.10	-0.03	-0.99	-0.08	-0.39	0.51
T178 H γ	-1.35	-1.59	-1.60	-0.04	-1.30	-1.24	-1.30	-1.18	0.58
W182 H δ 1	0.16	0.01	0.02	-0.24	-0.25	0.02	-0.10	-0.09	0.13
W182 H ζ 2	-0.17	-0.23	-0.20	-0.25	-0.28	-0.18	-0.23	-0.23	0.04
W182 H η 2	0	-0.08	-0.12	-0.21	-0.21	-0.08	-0.14	-0.14	0.06
W182 H ζ 3	-0.08	-0.21	-0.22	-0.22	-0.21	-0.12	-0.13	-0.19	0.05
W182 H ϵ 3	-0.21	-0.22	-0.22	-0.25	-0.16	-0.14	-0.10	-0.18	0.06
resolution		1.55	1.9	2.9	3.5	3.0	3.5		
method		x-ray	x-ray	x-ray	x-ray	Em	em		
rms (all)		0.19	0.21	0.38	0.24	0.34	0.30	0.28	0.08
R (all)		0.82	0.79	0.32	0.72	0.48	0.56	0.61	0.19
rms(arom/met)		0.12	0.11	0.37	0.22	0.22	0.24	0.21	0.10
r(arom/met)		0.94	0.95	0.33	0.82	0.80	0.75	0.77	0.23
rms(core, no H β)		0.11	0.10	0.16	0.23	0.19	0.25	0.17	0.06
r(core, no H β)		0.94	0.97	0.89	0.64	0.83	0.49	0.79	0.19
rms(structure selection)		0.24	0.25	0.51	0.27	0.36	0.37	0.24	0.10
r(structure selection)		0.66	0.63	0.00	0.55	0.27	0.37	0.70	0.25

Tab. 3.4 Comparison of experimental and calculated structural chemical shift of different structures of all-trans bacteriorhodopsin. Root mean square deviation (rms) and experimental correlation coefficient (r) were calculated using equations (3.1) and (3.2). The last two columns are the average (av.) and standard deviation (std.) of the six values calculated from crystallographic structures. The following chemical shifts are included in the different rms/r values: rms/r(all): all shifts of Tab. 3.3; rms/r(arom/met): all methyl (H γ) and aromatic (H δ 1, H ζ 2, H η 2, H ζ 3, H ϵ 3); rms/r(core, no H β): H α and H γ of T90 and T142 and aromatic protons of W86 and W182; rms/r(structure selection): shifts used for the selection of NMR families of structure (T90, T142, T178 and W86).

The ordering of chemical shift rms and resolution of different structures published by one single group is strictly fulfilled. Thus the chemical shift deviations get less in the refined electron microscopic structures 2at9 and 2brd compared to 1at9 and 1brd respectively. The same holds true comparing all-trans structures at lower resolution published later in time, the refinement of which is based on data of the same group at

higher resolution (rms 1c3w (resolution 1.55 Å) < rms 1c8r (resolution 1.8 Å) < rms 1brx (resolution 2.3 Å) for the three structures of Luecke et al.).

A closer inspection of the rms values between calculated and observed chemical shifts of the core residues (Thr 90, 142 H α and H γ , Trp 86, 182 H δ 1, H ζ 2, H η 2, H ζ 3, H ϵ 3) suggests that the deviations correlate with the resolution of the crystal structure. The structure with the highest resolution shows the lowest rms of the chemical shifts. To examine the significance of that correlation we chose the following approach: (1) the structure with the highest resolution (1c3w) is taken as standard for comparing different coordinates rms and chemical shift rms values. (2) The chemical shift rms for all all-trans coordinates found in the Brookhaven protein data base are calculated and the difference to the chemical shift rms to 1c3w is calculated. (3) The structures are superimposed on the backbone atoms of the core region of the seven helices of 1c3w and the rms of the side chain heavy atoms of Thr 90, Thr 142 and Trp 86, 182 are calculated.

A graphical representation of the obtained difference in side-chain heavy atom rms and chemical shift rms are shown in Fig. 3.5. From these data a linear correlation between the local accuracy of the structures and the chemical shift rms is evident.

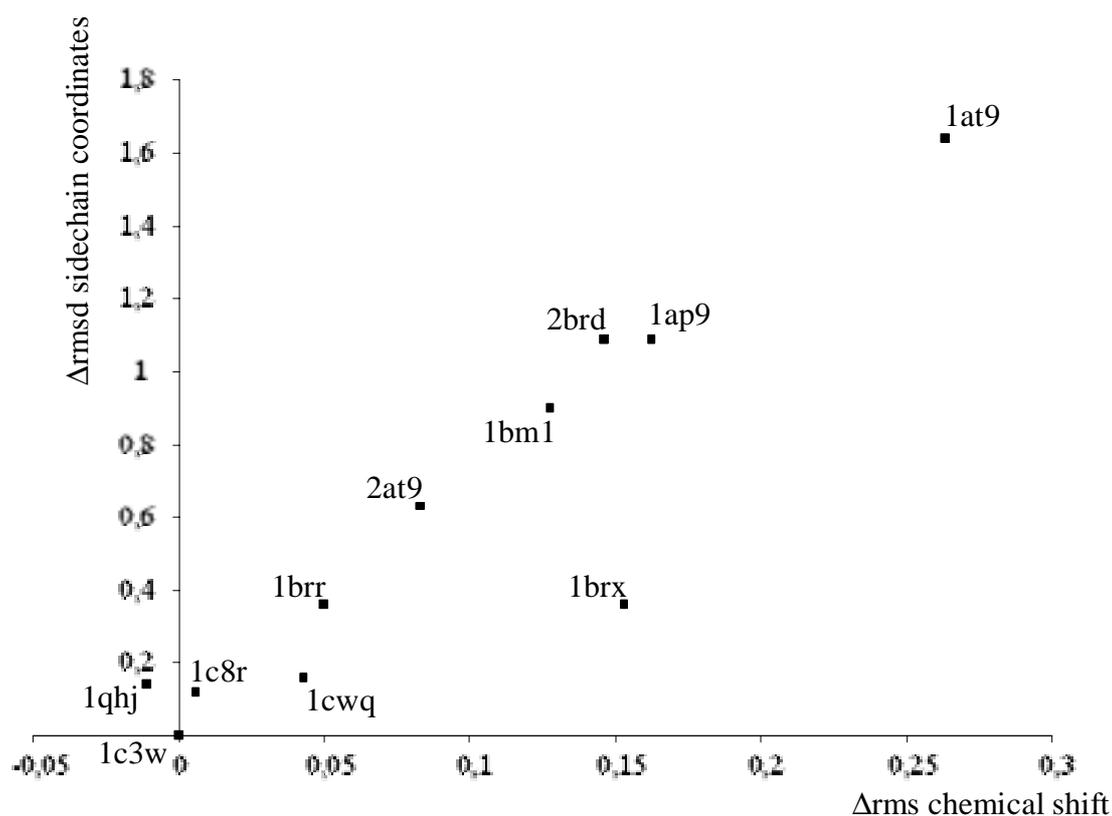


Fig. 3.5 Correlation between chemical shift difference and coordinate rms of the core residues (Thr 90, 142 without H β , Trp 86, 182) of different crystal structures. The labels correspond to the pdb entry code of the coordinates: 1c3w (1.55 Å) (Luecke, 1999b), 1c8r (1.8 Å) (Luecke, 1999a), 1qhj (1.9 Å) (Belrhali, 1999), 1cwq (2.25 Å) (Sass, 20-OCT-99), 1brx (2.3 Å) (Luecke, 1998), 1at9 (2.5 Å) (Pebay-Peyroula, 1997), 1brl (2.9 Å) molecule B of (Essen, 1998), 1bm1 (3.5 Å) (Sato, 1999), 2at9 (3.0 Å) (Mitsuoka, 1999), 1at9 (3.0 Å) (Kimura, 1997) 2brd (3.5 Å) (Grigorieff, 1996)

3.6 Derivation of distance constraints

3.6.1 Determination of a distance constraint list from NOE volumes

NOE cross-peak intensities were translated into distance constraints by either integration and comparison to a reference interaction, or estimated and used as a distance range in the calculations. Non-overlapped cross-peaks appearing in spectra with mixing times of 10 ms or less were integrated, assuming direct proportionality between the NOE intensity and the distance between the two contributing proton types. The average volume of four isolated intra-residue NOEs of Trp86 (H η 2/H ζ 2, H η 2/H ζ 3, H ζ 2/H η 2, H ζ 3/H η 2) was used as reference volume with a reference distance of $r_{\text{ref}}=2.485 \text{ \AA}$.

Distances were calculated by

$$r = r_{\text{ref}} \sqrt[6]{\frac{nI_{\text{ref}}}{I}} \quad (3.4)$$

where I is the integrated intensity, r_{ref} and I_{ref} are the reference distance and integral, respectively, and n is the number of protons involved in the NOE (Fletcher, 1996). Overlapping peaks in spectra with 5 and 10 ms were assigned to a distance range by comparing the isolated part of the peak volume to reference peaks.

Cross-peaks in spectra with mixing times of 20, 40, and 80 ms were assumed to represent a distance contained in 3 distance range classes, taking into account that spin diffusion might affect the peak intensity considerably. The distance range of class I was set to 3.0-4.5 \AA , of class II to 4.0-6.5 \AA , and of class III to 5.0-8.0 \AA . Cross-peaks were assigned to a class after comparison of their intensity with those of intra retinal NOEs with distances in the range of the boundary between two classes. Peaks used for comparison were (1) H15-H10, (2) H15-H20 and (3) H15-H16/17 NOEs of 13-cis (see Fig. 3.1). Peaks with intensities in between (1) and (2) were put in class I, those between (2) and (3) in class II and those weaker than (3) in class III. Practically, peaks in spectra with volumes too small to be integrated reliably in spectra with 10ms, but clearly above the noise level in spectra with 20ms mixing time were put into class I. Those which appear clearly above the noise in spectra with 40ms mixing time were

put into class II and all other into class III. This simple treatment is justified because the signal to noise ratios in the spectra of different samples used for distance calibration are similar. The specifications of the lower boundaries for the distance classes are to some extent arbitrary. To be on the safe side, the lower boundaries should be assigned to the van der Waals distance of the two atoms, because the intensity of a NOE is decreasing in the case of internal motions. From neutron scattering data it is known that the core of bacteriorhodopsin is very rigid and so we neglect internal motions as a source of significant error in the NOE volumes. The specific values for the lower boundaries in the distance calculations were carefully crosschecked in the evaluation of the refined structures. If NOE violations occurred at the lower boundary of a class, the boundary was relaxed and the calculation repeated.

A proper error analysis for peak volume integration in multi-dimensional NMR spectra is difficult because of the long measurement time of single spectra. This leads typically to cases, where only a very small number of spectra are recorded and a proper statistical analysis from the comparison of several independent measurements is not possible. In the bacteriorhodopsin spectra, additional systematic errors are caused by baseline distortion from the protonated detergent. The resulting errors are not randomly distributed.

To get an estimate for the relative differences between similar peaks, we did some simple statistical analysis on the four calibration peaks in all spectra with mixing times of 20 ms and shorter of samples containing ^1H Trp. A total of 9 spectra were analyzed, in which the peaks CIS H η 2 H ζ 2, H η 2 H ζ 3 and their diagonal counterparts were integrated. The peak volumes were normalized to the H ζ 2 H η 2 peak of each spectrum. The proton-proton distances H η 2-H ζ 2 and H η 2-H ζ 3 for tryptophan side-chains of 2.50 Å and 2.47 Å result in a NOE ratio H η 2-H ζ 2:H η 2-H ζ 3 of 1.002 using the linear approximation Eq. (3.4). Taking the peak volumes of the 9 different spectra with mixing times between 5 and 20 ms, we get a total of 20 ratios. The average of these is 1.11 and the estimated mean squared error or standard deviation of a single measurement (Bronstein, 1997) is 1.50. Thus we get a *relative error* of approximately 135% in the peak volumes. Since this error gives a physically unrealistic negative number as lower boundary for the confidence interval, we use the smaller most *probable error* of 91% in the peak volume. The real value is with a probability of 0.5

within the error limits specified by the most probable error (0.68 for the relative error). For the calibration distance, which is the average of four measured values, the error is half the error of a single measurement (the error of the mean scales with the square root of the number of measurements n). Taking the calibration distance $r_{\text{cal}}=2.485 \text{ \AA}$ without error (the relative error of the distance can be neglected compared to large error of the volumes), we can calculate the upper and lower bounds for the distance by plugging the upper and lower bounds into Eq. (3.4). The result is a confidence interval, where the lower bound is 5% less and the upper bound is 42% more than the actual result, e.g. for a NOE derived distance of 3 \AA , we get an confidence interval [2.87 \AA , 4.07 \AA]. These errors have to be taken as a relative rough estimate, since we have a small statistical sample and we neglect any systematic error.

During the structure calculation, the boundaries of the zero part in the NOE energy term were set to $\pm 10\%$ for the integrated peaks. The reasoning to use rather tight boundaries is to emphasis the NOE energy terms compared to the other energy terms in the force field. All distance constraints used for the calculations are summarized in Tab. 3.5.

	<i>13-cis, 15-syn</i>	<i>all-trans, 15-anti</i>
<i>intra-retinal</i>		
H7-C16H ₃ /C17H ₃	4.0-6.5	4.0-6.5
H7-C18H ₃	VDW-2.9	VDW-2.9
H7-C19H ₃	VDW-2.9	2.7±10%
H8-C16H ₃ /C17H ₃	VDW-2.9	VDW-2.9
H10-H15	3.0-4.5	-
H11-H15	4.0-6.5	4.0-6.5
H11-C19H ₃	2.6±10%	2.6±10%
H11-C20H ₃	2.7±10%	2.8±10%
H14-H15	3.0-4.5	3.0-4.5
H14-C20H ₃	VDW-2.9	4.0-6.5
H15-C20H ₃	4.0-6.5	VDW-2.9
H12-H15	2.1±10%	4.5-8.0
H15-C16H ₃ /C17H ₃	4.5-8.0	-
<i>retinal / tryptophan</i>		
H8-W189δ1	4.5-8.0	-
H12-W86δ1	3.6±10%	4.0-6.5
H14-W86δ1	4.0-6.5	VDW-3.5
H15-W86δ1	3.5±10%	-
C16H ₃ /C17H ₃ -W86ε3	3.0-4.5	4.0-6.5
C16H ₃ /C17H ₃ -W86ζ3	VDW-2.9	VDW-2.9
C16H ₃ /C17H ₃ -W86η2	3.0-4.5	4.0-6.5
C16H ₃ /C17H ₃ -W86ζ2	4.5-8.0	-
C16H ₃ /C17H ₃ -W189δ1	VDW-2.9	VDW-3.5
C19H ₃ -W182δ1	-	4.5-8.0
C20H ₃ -W86δ1	4.0-6.5	-
C20H ₃ -W182δ1	3.0-4.5	3.0-4.5
C20H ₃ -W182ε3/ζ3/η2/ζ2	4.0-6.5	3.0-4.5
<i>retinal / threonine</i>		
H10-T90γH ₃	4.0-6.5	-
H11-T90γH ₃	3.0-4.5	3.0-4.5
H14-T90γH ₃	4.0-6.5	-
H15-T90γH ₃	4.5-8.0	4.5-8.0
C19H ₃ -T90γH ₃	4.0-6.5	4.5-8.0
C20H ₃ -T90γH ₃	VDW-3.5	3.0-4.5
C18H ₃ -T142α	VDW-2.9	VDW-2.9
C18H ₃ -T142γH ₃	4.0-6.5	4.0-6.5
<i>threonine / tryptophan</i>		
T90γH ₃ -W86δ1	4.5-8.0	4.5-8.0
T90γH ₃ -W182δ1	4.0-6.5	4.5-8.0
T90γ-W182ε3/ζ3/η2/ζ2	4.0-6.5	3.0-4.5
T142γ-W138ε3/ζ3/η2/ζ2	4.0-6.5	4.0-6.5
T178γ-W182δ1	3.0-4.5	3.0-4.5
T178γ-W182ε3/ζ3/η2/ζ2	4.0-6.5	4.5-8.0

Tab. 3.5 Distance constraints used in the structure calculations.

Additional distances were derived for the pairs Thr67/Trp80 and Thr121/Trp137. The confidence intervals for these distances are taken as -5% and +42% of the determined values. The same confidence interval is specified for the intra threonine cross-peaks. However these peaks are in a region of the spectrum with additional noise due to residual protons of the deuterated dodecyl. Therefore additional systematic errors in

the peak volumes are expected. The distances are compared to different coordinate sets from the pdb. A comparison to the structures is made in chapter 3.7.4.

			crystal structures					
	NMR	confidence	1c3w	1qhj	1br	1bm1	2at9	2brd
T67-W80								
H α -H δ 1	3.89	3.70-5.52	4.02	4.22	4.49	5.16	4.17	11.3
H β -H δ 1	2.41	2.29-3.42	2.19	2.23	2.32	2.96	3.41	14.1
H γ -H δ 1	2.97	2.82-4.22	2.84	2.70	2.78	5.42	1.81	13.7
T121-W137								
H α -H ϵ 3*	2.72	2.58-3.86	3.24	3.26	3.74	3.20	3.71	4.59
H β -H ϵ 3*	2.21	2.10-3.14	2.09	2.31	4.58	4.57	2.05	4.78
H γ -H ϵ 3*	3.06	2.91-4.35	3.21	3.31	2.56	2.71	3.44	2.33
T67								
H α -H γ	3.38	3.21-4.80	3.81	3.79	3.78	3.11	3.56	3.10
T121								
H α -H γ	3.68	3.50-5.23	2.86	2.88	3.18	3.19	2.91	3.06
H β -H γ	2.80	2.66-3.98	2.50	2.48	2.45	2.45	2.45	2.46

Tab. 3.6 NMR derived distance constraints for two threonine tryptophan pairs outside the chromophore region. The * indicates that the specific assignment of this proton is based on the coordinates of 1c3w. The last six columns give the distances from the highest resolution crystal structures of six different research groups (1c3w (Luecke, 1999b), 1qhj (Belrhali, 1999), 1br molecule B of (Essen, 1998), 1bm1 (Sato, 1999), 2at9 (Mitsuoka, 1999), 2brd (Grigorieff, 1996)).

Many of the distance constraints of Tab. 3.5 and Tab. 3.6 are between residues of different helices. Retinal is connected to Lys 216 in helix G and shows NOEs to residues of helix C (which is formed by residues 36-64), helix E (133-155) and helix F (164-192). Additional NOEs are observed between helix D (104-127) and helix E, as well as between the small beta sheet region of the BC loop (67-71 and 74-78) and the beginning of helix C. If the methionine assignments of chapter 3.4.3 are confirmed, additional distance constraint are observed between helices A (9-31) and G (Trp12/Met209) and between helices D and F (Met118/Trp182).

3.6.2 Back-calculation of NOESY spectra

With the usual approximations made in the derivation of the master equation of the reduced density operator, the NOESY cross-peak volume after a mixing time τ_m is (e.g. (Borgias, 1990))

$$I_{i,j} = [\exp(-\mathbf{R}\tau_m)]_{i,j} M_{j,0} \quad (3.5)$$

where $M_{j,0}$ is the equilibrium magnetization of spin j . The relaxation matrix elements R_{ij} are functions of the transition rates between energy levels. For a system of n spins \mathbf{R} is a $(n \times n)$ matrix. For dipolar relaxation, the elements of \mathbf{R} are

$$R_{i,j} = \begin{cases} w_2^{i,j} - w_0^{i,j} & \text{if } i \neq j \\ \sum_{k \neq j} w_0^{k,j} + 2w_1^{k,j} + w_2^{k,j} + R_{ext} & \text{if } i = j \end{cases} \quad (3.6)$$

The diagonal elements are given by the longitudinal relaxation rates. They are determined by the dipolar interaction of spin j with all other spins in the sample plus an external rate, which takes care about other relaxation mechanisms. The cross-relaxation rates $R_{i,j}$ are given by the difference between the rates of simultaneous flips of spins i and j in the same sense (both spins make a transition from spin up to spin down or vice versa) and in an alternate sense (one spin makes a transition up/down, the other one down/up or vice versa). The transition rates w_2^{ij} (flip-flip term) and w_0^{ij} (flip-flop term) are determined by the dipolar interaction strength of the spins and the model for the random motion of the spins, which enters through a spectral density function.

$$\begin{aligned} w_0^{i,j} &= \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^4 \eta^2}{10 r_{i,j}^6} J(0) \\ w_1^{i,j} &= \frac{3}{2} \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^4 \eta^2}{10 r_{i,j}^6} J(\omega) \\ w_2^{i,j} &= 6 \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^4 \eta^2}{10 r_{i,j}^6} J(2\omega) \end{aligned} \quad (3.7)$$

where γ_I is the gyromagnetic ratio (, which is assumed to be the same for the two interacting spins) and $r_{i,j}$ is the distance between spins i and j . The spectral density

$J(\omega)$ is a function of the rotational correlation time τ_c . For protein structure determination an isotropic rotation of a rigid body it is often assumed, which leads to

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (3.8)$$

It is important to note that the NOE intensity (Eq. (3.5)) of a cross-peak between spin i and j depends on the complete relaxation matrix \mathbf{R} and thus on the distances to all spins in the sample. To get an analytical solution of the cross-peak intensities in terms of the transition rates and thus of the distances, the matrix \mathbf{R} has to be diagonalized and the Eigenvalue problem has to be solved. An analytical solution can in principle be obtained for up to five spins (Landy and Rao, 1989).

The matrix exponential can be approximated by a Taylor series in τ_m . For sufficiently short mixing times, terms of higher order can be neglected and the NOE intensities are given by

$$I_{i,j} = M_{j,0} [\delta_{i,j} - R_{i,j} \tau_m] \quad (3.9)$$

In this expression, the cross-peak intensity of the NOE between spin i and j depends on the term $R_{i,j}$ only, which is proportional to the inverse of the sixth power of the distance between the two spins. This is the approximation, which was made in the quantification of the peak intensities in chapter 3.6.1.

After the determination of the structure, numerical solutions of Eq. (3.5) can be used to calculate build-up rates of NOE cross-peak intensities (Keepers and James, 1984; Yip and Case, 1989). This approach has been used to refine protein structures. In the present case, the build-up curves for individual cross-peaks can't be determined accurately enough to perform such a refinement. However, from the calculated structures, we can back-calculate NOESY volumes to validate the initial rate approximation and to look at qualitative features of the build-up curves. All calculations have been performed on a coordinate file with the retinal in 13-cis, 15-syn configuration using the routine implemented in XPLOR (Brünger, 1992; Nilges, 1991).

We first address the question whether the linear approximation for the integrated cross-peaks is valid in the spectra with 10 ms mixing time. We calculate NOE

volumes from a coordinate file and evaluate the obtained data subsequently in the same way as the experimental data. Distances are calculated using Eq. (3.4), with a reference distance of $r=2.485\text{\AA}$ and a reference volume given by the average of the simulated volumes of the intra tryptophan cross-peaks W86 H η 2-H ζ 2 and W86 H η 2-H ζ 3. In the experimental data, the observation of a NOE in the spectra with short mixing times depends on the signal-to-noise ratio. Weak peaks cannot be integrated, because they are under the noise floor. This limitation does not apply to the simulated data and we analyze the peak volumes of all assigned NOEs.

An open parameter in the simulations is the rotational correlation time of the protein. In Appendix D we get an estimate for the rotational correlation time. To allow some variation, we performed calculations with different values. The resulting distances are listed below together with the ‘true’ distance of the corresponding protons in the coordinate file used during the simulation.

	Input distance [\AA]	Distance determined with linear approximation			
		$^2\text{H BR}$	$^1\text{H Ret}$	$^1\text{H Trp}$	$^1\text{H Thr}$
		$\tau_c=40\text{ns}$	$\tau_c=50\text{ns}$	$\tau_c=60\text{ns}$	$\tau_c=40\text{ns}$
<i>intra-retinal</i>					
H15-H14	3.11	3.09	3.08	3.08	3.13
H15-C20H ₃	4.63	4.26	4.19	4.13	4.09
H15-H12	2.11	2.18	2.19	2.21	2.19
H15-H11	4.71	4.05	3.97	3.90	3.82
H15-H10	4.02	3.09	3.02	2.97	2.92
H15-C16H ₃	7.91	6.71	6.44	6.21	5.99
H15-C17H ₃	7.60	6.72	6.49	6.28	5.87
H14-C20H ₃	2.80	2.77	2.76	2.76	2.78
H11-C19H ₃	2.73	2.75	2.76	2.76	2.69
H11-C20H ₃	2.69	2.72	2.73	2.74	2.65
H7-C19H ₃	2.62	2.67	2.68	2.70	2.64
H8-C16H ₃	2.61	2.70	2.72	2.74	2.69
H8-C17H ₃	2.66	2.76	2.78	2.80	2.72
H7-C18H ₃	2.53	2.58	2.59	2.60	2.57
H7-C16H ₃	4.44	4.29	4.26	4.24	4.14
H7-C17H ₃	4.46	4.33	4.31	4.29	4.09

<i>retinal / tryptophan</i>					
H15-W86δ1	2.98	2.96	2.96	2.96	2.89
H14-W86δ1	4.60	4.40	4.36	4.33	4.13
C20H ₃ -W86δ1	6.44	5.91	5.80	5.70	5.51
H12-W86δ1	3.46	3.35	3.33	3.31	3.21
C16H ₃ -W86ε3	3.07	3.07	3.07	3.07	3.00
C17H ₃ -W86ε3	4.40	3.97	3.90	3.84	3.73
C16H ₃ -W86ζ2	6.74	6.02	5.88	5.74	5.61
C17H ₃ -W86ζ2	5.44	4.98	4.89	4.81	4.61
C16H ₃ -W86ζ3	3.20	3.17	3.17	3.17	3.08
C17H ₃ -W86ζ3	2.88	2.92	2.93	2.94	2.85
C16H ₃ -W86η2	5.36	4.58	4.47	4.38	4.28
C17H ₃ -W86η2	3.72	3.61	3.58	3.56	3.48
C20H ₃ -W182δ1	3.63	3.56	3.55	3.54	3.39
C20H ₃ -W182ε3	6.59	6.12	6.01	5.91	5.79
C20H ₃ -W182ζ2	3.97	3.85	3.82	3.80	3.65
C20H ₃ -W182ζ3	7.10	6.45	6.31	6.17	6.06
C20H ₃ -W182η2	6.05	5.19	5.06	4.95	4.83
H8-W189δ1	6.60	5.01	4.87	4.74	4.61
C16H ₃ -W189δ1	5.47	4.76	4.65	4.56	4.43
C17H ₃ -W189δ1	3.59	3.51	3.49	3.48	3.41
<i>retinal / threonine</i>					
H15-T90γH ₃	6.46	5.73	5.62	5.51	5.39
H14-T90γH ₃	5.38	4.77	4.67	4.58	4.54
C20H ₃ -T90γH ₃	3.59	3.30	3.29	3.28	3.21
H11-T90γH ₃	4.00	3.93	3.91	3.90	3.75
H10-T90γH ₃	5.16	5.03	5.00	4.97	4.72
C19H ₃ -T90γH ₃	5.10	4.60	4.57	4.53	4.26
C18H ₃ -T142α	2.73	2.74	2.74	2.74	2.73
C18H ₃ -T142γH ₃	4.81	4.11	4.05	3.98	3.80
<i>threonine / tryptophan</i>					
T90γH ₃ -W182δ1	6.36	5.76	5.66	5.56	5.17
T90γH ₃ -W86δ1	6.71	6.26	6.16	6.07	5.62
T90γ-W182ε3	5.98	5.59	5.50	5.42	5.33
T90γ-W182ζ2	2.68	2.68	2.69	2.69	2.71
T90γ-W182ζ3	4.99	4.49	4.41	4.33	4.32
T90γ-W182η2	3.22	3.15	3.13	3.12	3.12
T178γ-W182δ1	3.67	3.56	3.54	3.51	3.44
T178γ-W182ε3	5.73	5.54	5.49	5.44	5.17
T178γ-W182ζ2	4.63	4.58	4.57	4.56	4.10
T178γ-W182ζ3	6.63	6.30	6.22	6.15	5.86
T178γ-W182η2	6.19	5.84	5.77	5.71	5.25

Tab. 3.7 Comparison of distances determined from simulated NOE intensities using the linear approximation. The simulations were performed for two different labelling patterns and three different mixing times. The first column of distances corresponds to the distance of the spins in the input structure file for the simulation. Note that these distances are taken from a preliminary 13-cis, 15-syn structure and do not correspond exactly to those of a member of the final family of structures.

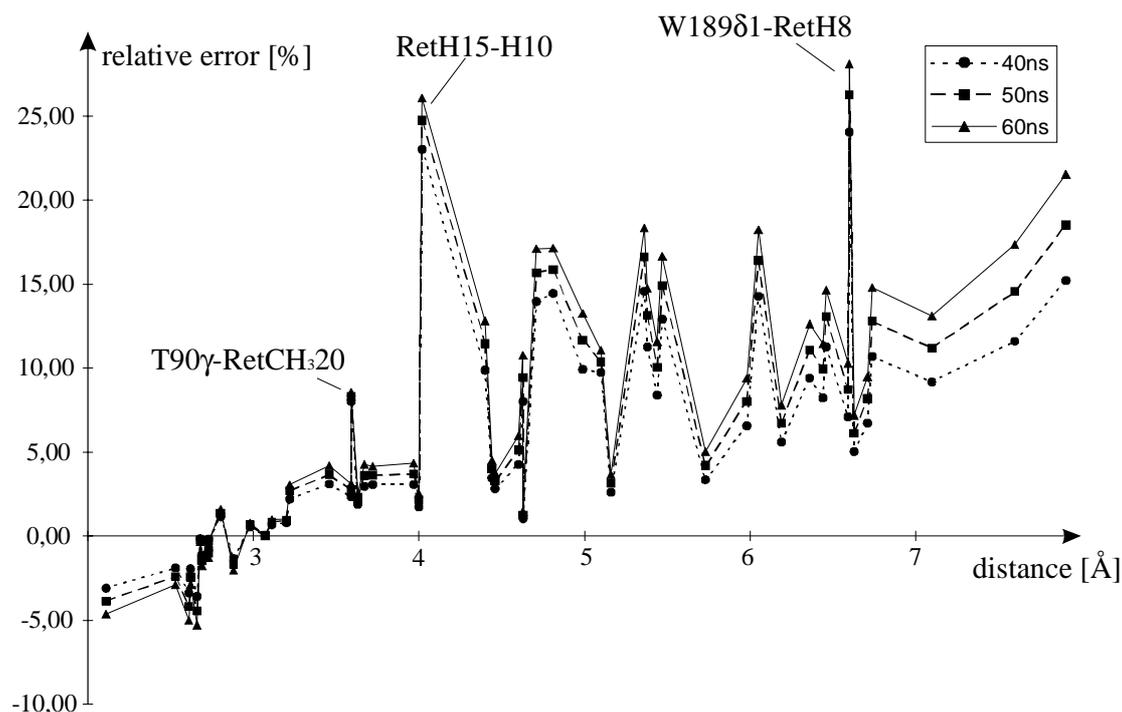


Fig. 3.6 Error of the distances determined by the linear approximation relative to the true distances in the structure used for the calculation of the NOE by the relaxation matrix approach. The relative error is $100 \cdot (r_{\text{true}} - r_{\text{linear}}) / r_{\text{true}}$. A positive error thus corresponds to a distance shorter than the true distance.

The graphical representation of the relative error induced by the linear approximation Fig. 3.6 shows that the different curves are qualitatively the same. The effect of longer correlation times is merely to shift the region of applicability of the initial rate approximation to shorter mixing times (Madrid, 1989). Different τ_c values can be related to each other by an appropriate scaling of t_m (Post, 1990). The relative error introduced by the linear approximation and the way of calibration is less than $\pm 5\%$ for all distances up to 3.5 \AA . Peaks between protons further apart than 3.5 \AA are typically close to the noise level in the experimental spectra. So we can conclude, that errors by the linear approximation is less than 5% for the integrated peaks in spectra with mixing times up to 10 ms.

The second question is, whether the sample labelling causes specific errors in the NOE volumes, which are not observed in uniformly ^1H labelled samples. To test this, two simulations were performed with different labelling patterns, but the same mixing time and rotational correlation time. Fig. 3.7 is a graphical representation of the

relative errors. Again, both data sets have the same general features. However, the curve cannot be superimposed by simple scaling.

In general, the errors caused by the linear approximation are reduced in the selectively labelled sample (see Tab. 3.7). Exceptions are some of the distances, which are shorter than the calibration distance. It is interesting to note that selective labelling has qualitatively the same effect than reducing the rotational correlation time for longer distances.

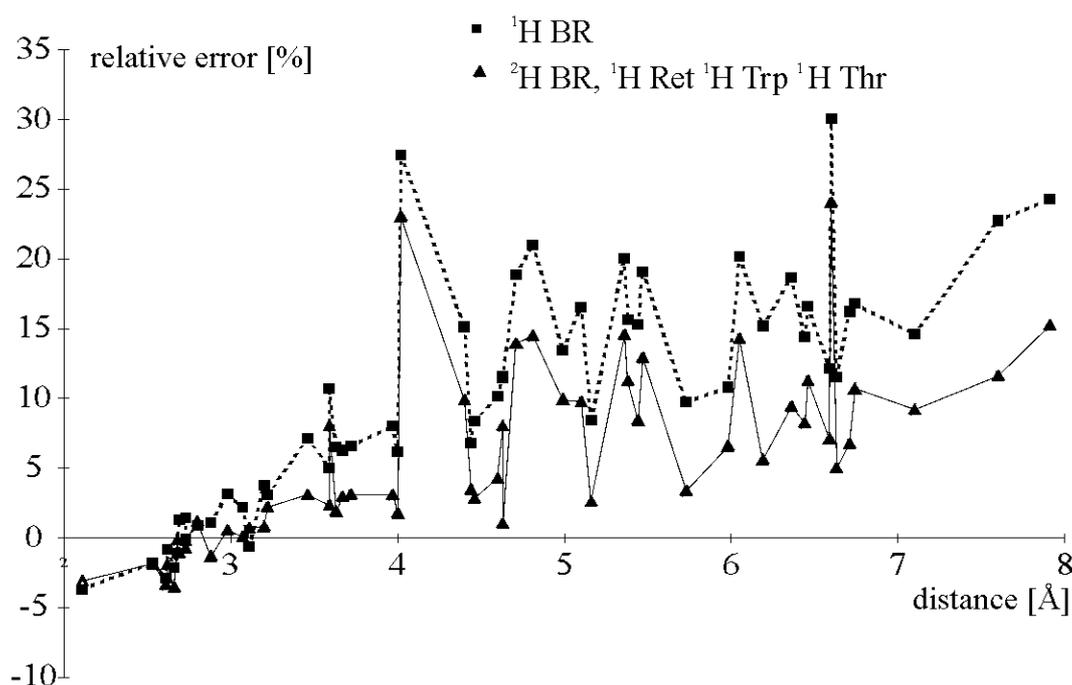


Fig. 3.7 Comparison of the relative errors introduced by the linear approximation in uniformly and selectively protonated bacteriorhodopsin.

The errors introduced by the linear treatment have three different sources:

1. Even in the case of an isolated spin pair, the linear treatment is not correct. The error leads to an overestimation of the very short distances. It shows a zero crossing at the calibration distance and levels off to a constant value. The value of this constant error for long distances scales with the mixing time.
2. The biggest errors due to spin diffusion are observed in a linear arrangement for the NOE between the outer spins. The volume of the cross-peak in this case is mainly determined by the indirect magnetization transfer through the intermediate spin. The distance between the outer spins is estimated shorter than it is in reality.
3. The NOE volume is reduced, if the longitudinal relaxation rate of the spins is getting faster. The leakage of magnetization can be caused by other protons nearby. This process will lead to errors in the estimation of distances, if the leakage rate is different for different spins.

The dependence of the relative error on distance in Fig. 3.6 and Fig. 3.7 is fully described by these three effects. The second and third point depend on the actual geometry and differ between different cross-peaks. This causes the ‘fine structure’ or ‘peaks’ on top of a general trend-line, which is defined by the error for two isolated spins. This general trend-line is large and negative for very short distances, shows a zero crossing at the calibration distance and levels off to a small error of less than 1% at large distances (the asymptotic value has been determined from calculation for a two-spin system with the same parameters as above).

The two distances with the largest relative errors (Ret H15-H10 and W189 δ 1-RetH8) can be explained by looking at the structure: Both NOEs are to a large extent mediated by a third spin, geometrically located in between the two spins forming the NOE. This is retinal H12 in the case of the retinal H15-H10 NOE and the two retinal methyls H16 and H17 in the case of the Trp189 δ 1 - retinal H8 NOE. The NOE intensity is mainly determined by the shorter distances to the third spin.

The third ‘peak’ annotated in Fig. 3.6 between Thr90 CH₃ γ and Ret CH₃20 corresponds to a difference in leakage rate. Both spins show strong NOEs to third spins. This leads to a faster leakage of magnetization and the error is getting bigger.

The leakage of magnetization is most obvious, if NOE intensities of the same cross-peaks in spectra with different labelling patterns are compared. The H ζ 3 proton of W86 is very close to the methyl groups 16 and 17 of the retinal. This leads to a much faster reduction of the W86 η 2 ζ 3 NOE at longer mixing times as compared to the W86 η 2 ζ 2 NOE. In spectra with deuterated retinal, the two peaks have almost the same intensities, as expected from the similar distances between the involved protons. Simulated buildup curves and two spectra recorded with the same mixing time are shown in Fig. 3.8. The initial rates for all four curves are approximately the same. The leakage rates in the absence of protonated retinal are slightly different, because the H ζ 2 proton is close to both H η 2 and H ϵ 3, while H ζ 3 is more isolated. The leakage increases for both protons in the presence of protonated retinal.

The example shows that by comparing intensities of identical peaks present in spectra with different labelling patterns, additional information about the local environment of

the spins is obtained. In the present case, the observation confirms the specific assignment of the Trp 86 AMXP spin system.

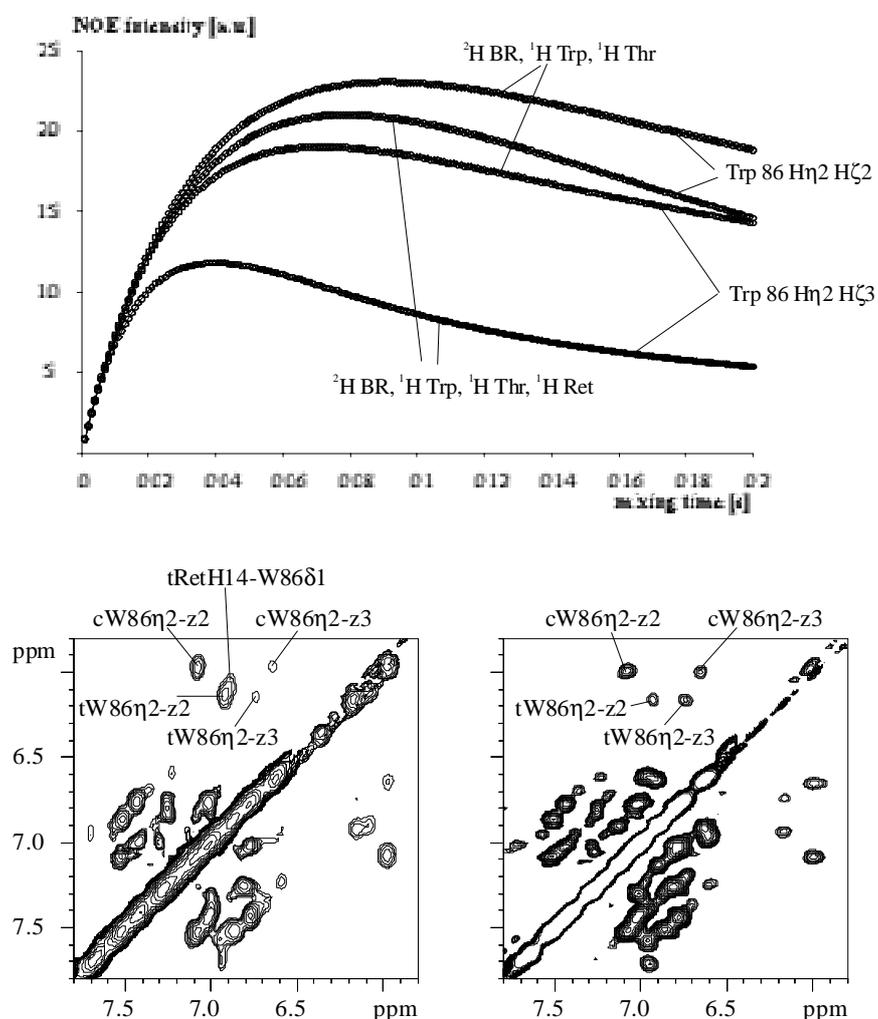


Fig. 3.8 NOESY spectra with mixing times of 40 ms. The sample on the left is $^1\text{H Ret}$, $^1\text{H Trp}$ and $^1\text{H Thr}$ labelled, the sample on the right is $^1\text{H Trp}$ and $^1\text{H Thr}$ labelled. The intensity of the cross-peaks between W86 H η 2 and W86 H ζ 3 is lowered in the Ret ^1H sample due to relaxation of the H ζ 3 proton by the close methyl groups CH₃ 16 and CH₃ 17.

3.7 Bacteriorhodopsin structures

3.7.1 Structure calculation

The NMR structures were calculated using the program AMBER (version 4.1) (Cornell, 1995), and the X-ray structure 1br (Essen, 1998) as a template. A simulated annealing protocol was applied in which only atoms within a distance of 10 Å from the retinal were allowed to move (elevating temperature to 1500 K at 0-5 ps, cooling down to 0 K at 5-10 ps with slowly increasing VDW repulsion terms from 0.001 to 1.0 between 0 and 5 ps) (Cornell, 1995). Point charges for a protonated chromophore have been calculated with GAMESS U.S., and the RESP programs following the method of Cieplak et al. (Cieplak, 1995; Schmid, 1993). Torsion barriers for the alternating single and double bonds of protonated retinal were set to 10 kcal/mol, according to quantum chemical calculations (Zhou, 1993). Asp 96 and Asp 115 were included as protonated (neutral) amino acids. Four hydrogen-bonded water molecules observed in the crystal structure of Luecke et al. (Luecke, 1998), 1brx) were included in the calculation. In the absence of the water molecule between Asp 85 and the Schiff base proton the structure calculation lead to violations of intra retinal NOEs. Hydrogen bonds in helices were restrained to their original values at the end of the calculation. Forces representing distance restraints were applied by way of a square-well function ($r_k=50.0 \text{ kcal.mol}^{-1} \text{ \AA}^{-2}$). A distance of 0.4 Å was added to a determined value to allow for methyl pseudo atoms, 1.5 Å was added to take account of the C16/C17 pseudo atom (Fletcher, 1996) and 2.4 Å was added for the pseudo atom representing Hε3, Hζ3, Hη2 and Hζ2 of a Trp side-chain.

For both all-trans and 13-cis,15-syn a total of 100 simulated annealing runs were performed. The final energy of the AMBER force field was very low for all calculated structures, with an average of -4944 ± 19 kcal/mol for all-trans structures and -4956 ± 18 kcal/mol for 13-cis, 15-syn structures. The number of NOEs is sufficient to get the correct conformation of the retinal in all calculated structures. During the slow cooling phase in the last 5 ps of the simulated annealing protocol, the force field energy is minimized effectively, leading to a low energy for all structures. In test calculations without NOE constraints, the retinal was found in different

conformations. The three structures with the lowest force field energy had the retinal in a 13-cis, 15-syn conformation.

The spread of the 100 calculated structures with the distance constraints of Tab. 3.5 defines the retinal position in the all-trans and the 13-cis forms, the position of the tryptophan side chains 86, 138, 182 and 189, and of threonines 90, 142, and 178. The largest variations are observed in the position of Trp 138, where two subgroups with different side-chain conformations are observed, leading to a rms of 0.42 ± 0.56 Å for the heavy side-chain atoms of this residue. (Tab. 3.8). Out of the family of 100 calculated structures - which we refer to as the *parent family* in what follows - 12 structures were selected based on minimum chemical shift rms (*chemical shift subfamily*) between the chemical shifts calculated from the coordinates and the observed values for the protons resonances of Thr 90, 142 and 178 ($H\alpha$, $H\beta$ and $H\gamma$) and Trp 86 ($H\epsilon_3, H\zeta_3, H\eta_2, H\zeta_2$ and $H\delta_1$); (14 chemical shifts in total). Trp 182 protons were not included, because their specific assignments are not uniquely determined from the spectra. The chemical shift of the protons of Trp 182 are very close to their random coil values and they barely vary. Including the shifts of Trp 182 did not lead to a significant difference in the selected structures. For 13-cis, 15-syn exactly the same 12 structures were selected if Trp 182 was included or not, for the all-trans family 11 out of 12 structures were identical.

The potential of the chemical shift rms of the assigned proton resonances for judging the quality of different bacteriorhodopsin structures is demonstrated in paragraph 3.5. However, since this usually is not applied, we make a more detailed comparison of the parent family, the chemical shift subfamily and a family of 12 structures with the lowest AMBER force field energies (*energy subfamily*). The average AMBER energies, average chemical shift rms and different coordinate rms values of both families of 12 selected all-trans structures and of all 100 calculated all-trans structures are compared in Tab. 3.8.

The first question addressed is: Does the selection of subfamilies lead to an improvement of the quality of the final model? In standard NMR structure determination, a selection of low energy structures certainly does improve the model because many of the calculated structures are trapped in completely erroneous three-

dimensional folds. In the present calculation, we start from a well-defined structure and keep the three-dimensional arrangement fixed. All structures converge to a low final energy. The subfamilies have of course lower average values of the selection criteria in comparison to the parent family. However, the average energy is larger in the chemical shift family and the average chemical shift rms is larger in the energy subfamily compared to the values of the parent family. Thus, the two selection criteria lead to different subfamilies and the chemical shift rms is not correlated with the AMBER energy. The backbone rms for the core of the protein (bb (sub)) and of all backbone atoms allowed to move during structure calculation (bb (heat)) do not change significantly for the parent family and the two subfamilies. The same is true for the side-chain of Trp 189. The side-chains of Trp 86 and Trp 182 show smaller rms values in the subfamilies, the lower values for the energy subfamily being significant. The most significant improvement is observed for Trp 138, where both criteria select structures of the same population of side-chain conformations. For the side-chain of Lys 216 and to a lower extent for the retinal carbon atoms, both of the subfamilies show considerably lower average rms values than the parent family.

The manifold of calculated structures defines the precision of the NMR structure. Both of the selection criteria lead to subfamilies, which are more precise in the positions of Trp 138 and Lys 216 as well as the retinal. The energy subfamily is more precise than the chemical shift subfamily. However, the increase in precision does not imply necessarily that the structures are closer to the true structure. To relate precision and accuracy statistically, specific assumptions have to be made for the probability distributions. Since such assumptions are very difficult to verify, the best way to get an estimate for the accuracy is to have an independent standard (Zhao and Jardetzky, 1994). Therefore, we compare both subfamilies to the crystal structure at highest crystallographic resolution (1c3w (Luecke, 1999b)), which is assumed to be closest to the true structure. The hypothesis for the high accuracy of that structure is supported by its chemical shift rms being smaller than for any of the structures in the parent family. In what follows the term *precision* is used as a synonym for the average rms of a family of structures with respect to its mean structure and the term *accuracy* as the average rms of a family of structures with respect to the crystal structure 1c3w.

The overlay of the two subfamilies and the crystal structure 1c3w in Fig. 3.9 immediately shows that the selected conformation of the Trp 138 side-chain is the correct one. Both criteria select more accurate structures than the parent family in this case. The accuracy of the Lys 216 side-chain in the chemical shift subfamily is with 1.02 ± 0.17 Å significantly higher than for the energy selected, the rms of which (1.17 ± 0.21 Å) is similar to the parent family (1.15 ± 0.20 Å). The lower accuracy in the Lys216 side-chain results in a significant lower rms of the retinal carbon atoms of the chemical shift subfamily as well. For the Trp side-chains other than Trp 138, the parent family and both sub-families show similar accuracy, except for Trp 182, which is slightly more accurate in the shift selected subfamily.

From these observations, we can conclude that although the energy selection leads to a higher precision of the structures, the accuracy does only improve significantly for Trp 138 and to a lower extend for Trp 86. For both of these residues, the parent family includes structures with very high accuracy and the energy selection does identify those structures successfully. However, for the Lys 216 side-chain and the retinal, the same is **not** true. Although, the energy selection does select a considerably more precise subfamily, the accuracy of this subfamily is even slightly worse than that of the parent family. The chemical shift criterion on the other hand selects structures which are both, significantly more precise *and* accurate.

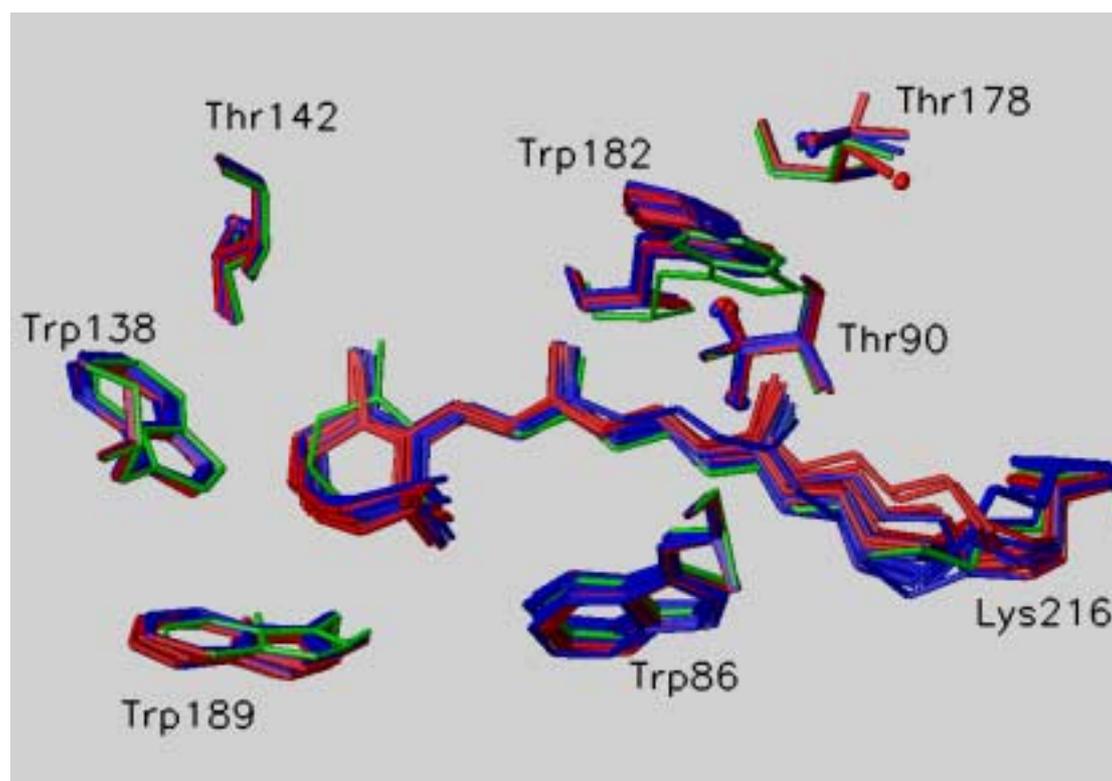


Fig. 3.9 Comparison of twelve all-trans structures selected out of 100 calculated simulated annealing runs with lowest AMBER energies (red) and with lowest chemical shift rms (blue). For comparison, the coordinates 1c3w are shown as well (green). The structures have been superimposed on to 1c3w (used as a reference structure) minimizing the root mean squared deviation of the backbone atoms (using MOLMOL (Konradi, 1996)) of residues 9-29 (helix A), 43-60 (helix B), 81-94 (helix C), 108-128 (helix D), 134-152 (helix E), 178-191 (helix F) and 205-221 (helix G). These residues were chosen for alignment as they constitute the central core of the helices in accord with (Subramaniam, 1999).

<i>number of structures</i>	<i>energy selected</i>	<i>chem. shifts selected</i>	<i>parent family</i>
<i>EAMBER</i>	12 -4970±7	12 -4935±20	100 -4944±19
<i>rms cs</i>	0.387±0.095	0.294±0.012	0.361±0.073
<i>rms relative to average structure</i>			
<i>bb (heat)</i>	0.08±0.01	0.09±0.01	0.09±0.02
<i>bb (sub)</i>	0.07±0.01	0.08±0.01	0.08±0.02
<i>Ret</i>	0.29±0.07	0.28±0.09	0.34±0.16
<i>Lys 216</i>	0.57±0.19	0.63±0.18	0.87±0.34
<i>Trp 86</i>	0.20±0.09	0.23±0.10	0.24±0.12
<i>Trp 138</i>	0.10±0.03	0.13±0.05	0.42±0.56
<i>Trp 182</i>	0.22±0.10	0.25±0.13	0.28±0.10
<i>Trp 189</i>	0.20±0.09	0.20±0.08	0.19±0.10
<i>rms relative to 1c3w</i>			
<i>bb (heat)</i>	0.32±0.01	0.32±0.01	0.32±0.01
<i>bb (sub)</i>	0.33±0.00	0.33±0.01	0.33±0.01
<i>Ret</i>	0.65±0.10	0.54±0.08	0.63±0.15
<i>Lys 216</i>	1.17±0.21	1.02±0.17	1.15±0.20
<i>Trp 86</i>	0.36±0.05	0.39±0.09	0.40±0.09
<i>Trp 138</i>	0.44±0.05	0.44±0.08	0.63±0.61
<i>Trp 182</i>	1.27±0.17	1.16±0.23	1.27±0.23
<i>Trp 189</i>	0.59±0.18	0.62±0.19	0.65±0.16

Tab. 3.8 Comparison of the statistics of twelve structures selected by either minimal force field energy (EAMBER) or minimal deviation of observed and calculated chemical shifts (rms cs) out of on hundred calculated coordinate files (column all). For the calculation of the coordinate root mean square deviation (rms), the coordinate files have been superimposed as in Fig. 3.9 on to either the average structure of the corresponding family (upper part) or on to 1c3w (Luecke, 1999b). The different coordinate root mean squared deviations include the following atoms: bb (sub) – backbone atoms of the core residues of the seven helices (Subramaniam, 1999), bb (heat) – backbone atoms, which are not fixed during the structure calculation, Ret – all retinal carbon atoms, Trp, Lys – all heavy side-chain atoms (including Schiff-base nitrogen for Lys 216).

To check whether this inaccuracy of Trp 182 is caused by an error in the distance constraints, the calculation was repeated with all constraints involving H δ 1 of Trp 182 removed. The assignment and distance determination of these cross-peaks is ambiguous since the H δ 1 and H ζ 2 resonances of Trp 182 are overlapping. The resulting parent family has a slightly higher average energy of -4936 ± 24 kcal/mol and a significantly higher average shift rms of 0.44 ± 0.12 compared to the parent family based on the full constraint list. The changes are caused mainly by a strong disorder in the Trp182 side-chain. The energy and chemical shift sub-families are shown in Fig. 3.10 together with the aromatic side-chains of Tyr185 and Phe208. Remarkably, the side-chains of these two aromatic amino acids show a very high precision, although there are no distance constraints in the structure calculation. This is in opposite to Trp 182, which is much less accurate for both subfamilies. The average rms of the Trp 182 side-chain to the X-ray structure is 1.17 ± 0.59 Å for the shift subfamily and 1.93 ± 0.94

Å for the energy subfamily. The rms between the mean structures of the chemical shift subfamilies of the two different calculations is 0.28 Å compared to a rms of 1.15 Å (full constraint average) and 1.02 Å (reduced constraint average) relative to the X-ray structure. It can therefore be concluded, that the reduced accuracy of the Trp 182 side-chain in the structure calculations is not caused by an error in the distance constraint file.

In summary we conclude that the chemical shift rms is a more useful parameter for the selection of a subfamily of structures than the AMBER force field energy in our calculations. The chemical shifts are independent experimental parameters, which are not used during the structure calculation. The selection based on chemical shifts therefore leads to a considerable improvement of the accuracy of the selected structures, especially for parts of the structure, which are less well defined. The force field energy on the other hand is minimized during the structure calculation. If it is used as a selection criterion as well, it can sort out structures, which are trapped in a higher energy state during the last part of the structure calculation, as for example the structures with alternative side-chain conformation of Trp 138. After having sorted out structures trapped in local minima, the remaining structures are distributed around the global minimum of the force field. If this minimum is inaccurate for some parts of the structure, only independent parameters - as chemical shifts for example - can be used to improve the final model.

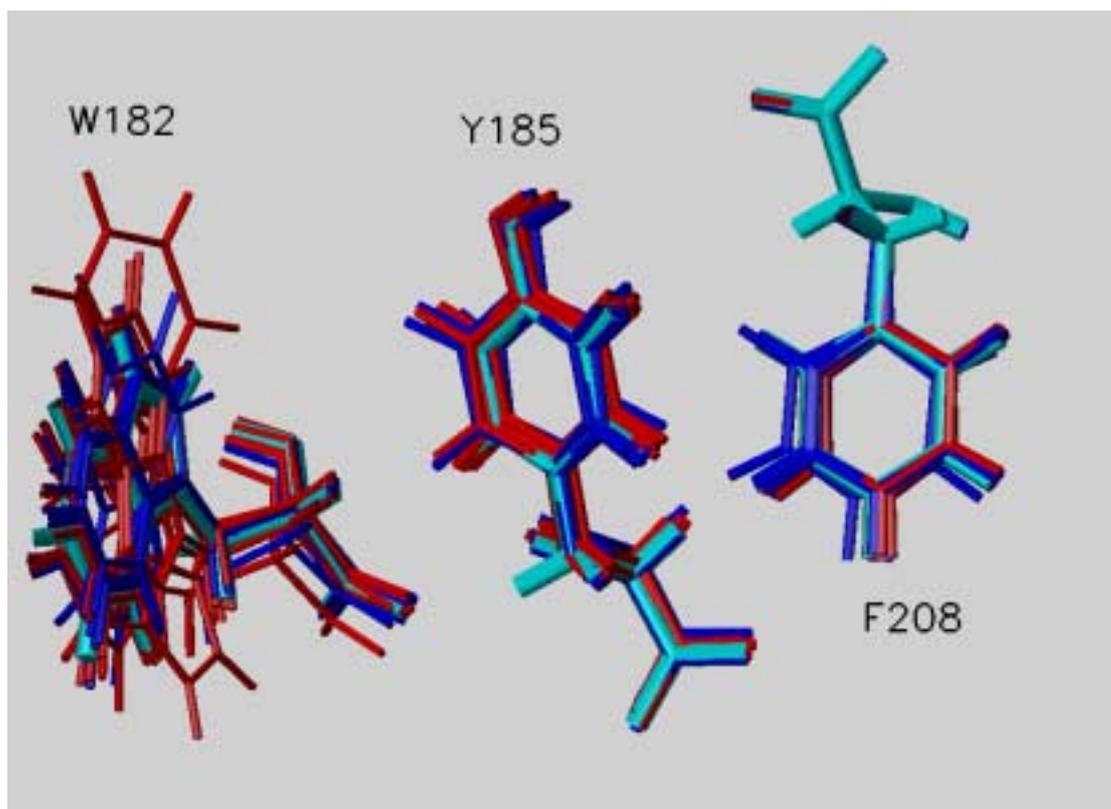


Fig. 3.10 Comparison of the side-chain orientation of three aromatic residues in structure calculations with modified constraints list (all constraints for W182 H δ 1 removed). The 12 structure of the energy subfamily (red) show much stronger variations in the side-chain orientation than the chemical shift subfamily (blue). The average structure of the chemical shift subfamily calculated with the full constraint lists is shown in cyan.

3.7.2 Comparison of the solution structures of the all-trans, 15-anti and 13-cis, 15-syn forms of bacteriorhodopsin

Comparing the two families of structures with the lowest chemical shift rms for the all-trans, 15-anti and 13-cis, 15-syn forms reveals the high similarity of the position of the retinal and the amino acid side-chains in the active site (Fig. 3.11). High similarity can be anticipated already from the small difference of the chemical shifts of similar protons in both conformations. The average rms of the tryptophan side-chains relative to the mean structure of the other conformer is only insignificantly higher than the average rms to the mean structure of the same conformer (Tab. 3.9). Essential difference is observed for the rms of the retinal carbon atoms. The largest deviations occur in the area around the Schiff-base where the geometry is different. The Schiff-base moiety is closer to the helix backbone and the methyl group attached to C20 moves slightly closer to Trp 182. The average distance between the protonated Schiff-base nitrogen and the two aspartic acids 85 and 212, which are part of the complex counterion, is found to be reduced in the 13-cis, 15-syn structure compared to the all-trans, 15-anti. The average distances between Lys 216 N ζ and Asp 85 O δ 2 are 4.53 ± 0.49 Å (all-trans) and 4.13 ± 0.51 Å (13-cis, 15-syn), respectively, and between Lys 216 N ζ and Asp 212 O δ 1 are 3.93 ± 0.67 Å (all-trans) and 3.44 ± 0.40 Å (13-cis, 15-syn), respectively. The significant reduction is caused by the retinal isomerization and changes in the Lys 216 sidechain conformation. The sidechains of Asp 85 and 212 do not show any significant changes between the two families of structure.

The subtle movement of the Schiff-base moiety is translated into a small displacement of the retinal carbon atoms C1-C11. The average rms of these atoms with respect to the mean structure is 0.25 ± 0.10 Å for both families of structure, whereas the average rms between one family to the average of the other is 0.38 ± 0.17 Å (see Tab. 3.9)

To give a more accurate quantification of the displacement of the atoms C1-C11 of the retinal in the all-trans, 15anti and 13-cis 15-syn families of structures, the following protocol was applied using the computer program XPLOR (Brünger, 1992):

1. All members of the all-trans family of structures were one after another rotated and translated into a coordinate system corresponding to the principle axis systems of the atoms C5-C15 of the retinal polyene chain. The x, y and z-axis correspond to the directions of the principle axis of inertia of the selected atoms with decreasing size of the principle moment of inertia. Thus, the x-axis points along the axis of inertia connected to the largest principle moment, which would be along a vector pointing in the long axis of a completely planar polyene chain. The y-axis would be perpendicular to x-axis in the plane of the polyene chain and the z-axis perpendicular to the plane of the polyene chain. Using internal coordinates, x corresponds approximately to the vector C15-C5 and y to C11-H11.
2. Each single 13-cis, 15-syn structure is fit to each of the oriented all-trans structures in order to minimize the backbone rms of the two structures.
3. For the two structures, the retinal atoms C1-C11 of the 13-cis, 15-syn structure are fit to those of the all-trans structure in order to minimize their rms. The components of the translation vector needed to achieve this fit are stored.
4. The average and standard deviation of the components for the translation vectors obtained by a pairwise fit of all 13-cis, 15-syn to all all-trans structures is calculated.

The result of the fitting procedure gives the pairwise average displacement of the retinal polyene chains together with a standard deviation, which can be interpreted as error in the estimate. The values are $x = -0.25 \pm 0.25 \text{ \AA}$, $y = 0.09 \pm 0.14 \text{ \AA}$ and $z = 0.06 \pm 0.13 \text{ \AA}$ for the components of displacement of 13-cis, 15-syn to all-trans. The average pairwise rms for these atoms is $0.45 \pm 0.18 \text{ \AA}$. To get an estimate of the significance of the above numbers, the same procedure was applied to all members of the all-trans and 13-cis, 15-syn families separately, resulting in values of $x = -0.06 \pm 0.28 \text{ \AA}$, $y = -0.02 \pm 0.13 \text{ \AA}$, $z = -0.04 \pm 0.10 \text{ \AA}$ and $\text{rms} = 0.37 \pm 0.16 \text{ \AA}$ for the 13-cis, 15-syn family and $x = -0.02 \pm 0.22 \text{ \AA}$, $y = -0.00 \pm 0.16 \text{ \AA}$, $z = -0.03 \pm 0.17 \text{ \AA}$ and $0.37 \pm 0.14 \text{ \AA}$ for the all-trans family.

We can conclude, that the polyene chain starting from C11 towards the β -ionone ring and the β -ionone ring of the retinal in the 13-cis, 15-syn structure is displaced by $0.25 \pm 0.25 \text{ \AA}$ along the direction of the polyene chain toward the Schiff-base relative to the same part of the all-trans structure. Perpendicular displacements are smaller than 0.1 \AA on average. The precision of the retinal structure as judged from the pairwise rms for the members of each family of structures is lower than the

determined displacement, which is reflected in an estimated error of the same order as the determined displacement.

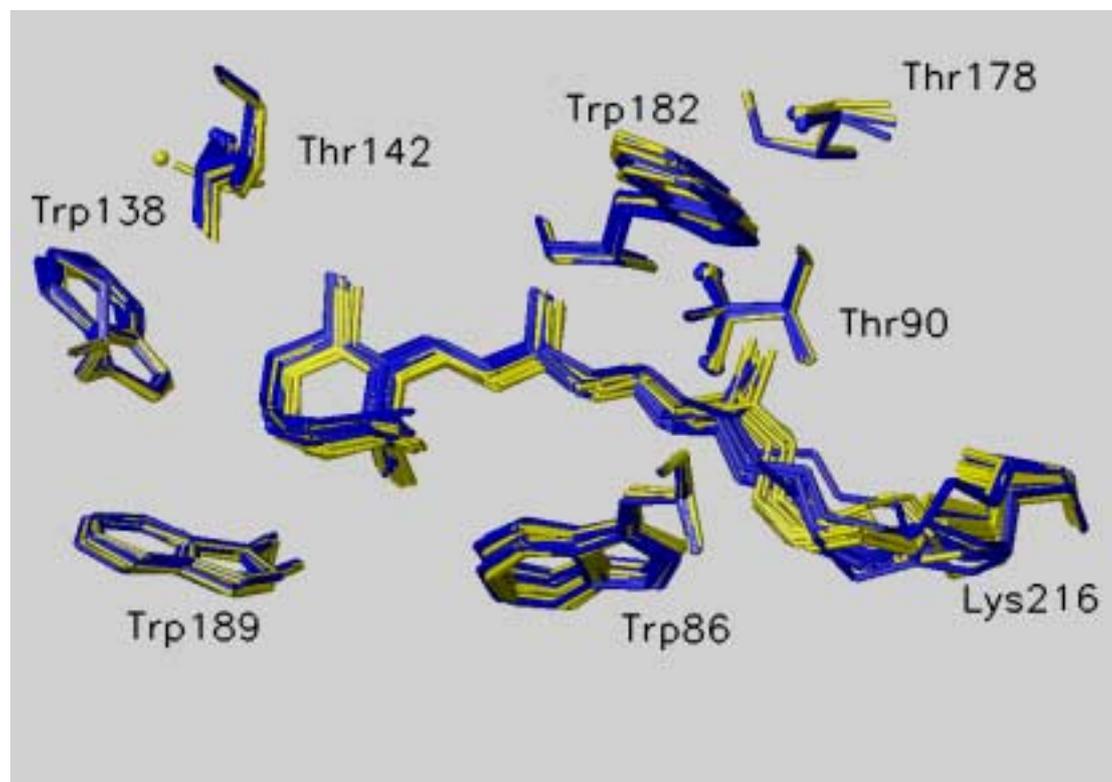


Fig. 3.11 Superposition of 12 calculated structures each of the *all-trans* (blue) and 13-*cis*, 15-*syn* forms (yellow) of detergent solubilized bacteriorhodopsin. The pseudo atoms representing the γ -methyl protons of the threonine side-chain are shown as spheres to allow distinguishing C γ and O γ .

	<i>13-cis, 15-syn family</i>	<i>all-trans, 15-anti family</i>
<i>Chemical shift rms</i>	0.282±0.007	0.294±0.012
<i>Amber energy</i>	-4963±17	-4935±20
<i>rms relative to the average structure</i>		
bb (sub)	0.07±0.02	0.08±0.01
Ret (all C)	0.26±0.09	0.28±0.09
Ret (C1-C11)	0.25±0.10	0.25±0.10
Lys 216	0.51±0.24	0.63±0.18
Trp 86	0.24±0.09	0.23±0.10
Trp 138	0.11±0.04	0.13±0.05
Trp 182	0.33±0.10	0.25±0.13
Trp 189	0.18±0.10	0.20±0.08
<i>rms relative to the average of the other family</i>		
bb (sub)	0.09±0.02	0.09±0.01
Ret (all C)	0.52±0.10	0.53±0.08
Ret (C1-C11)	0.38±0.17	0.38±0.14
Lys 216	0.80±0.09	0.86±0.16
Trp 86	0.28±0.09	0.26±0.12
Trp 138	0.12±0.07	0.14±0.07
Trp 182	0.35±0.15	0.30±0.14
Trp 189	0.19±0.05	0.21±0.09

Tab. 3.9 Statistics of the 12 selected structures of all trans and 13-cis, 15-syn forms. The rms values are determined as in Tab. 3.8.

Proton	13-cis, 15-syn family			all-trans family		
	exp.	clac.	calc.-exp.	exp.	calc.	calc.-exp..
W86 Hδ1	6.64	6.89±0.02	0.15	6.88	6.90±0.03	0.02
W86 Hζ2	7.06	6.87±0.09	-0.19	6.91	6.83±0.07	-0.08
W86 Hη2	5.97	5.89±0.26	-0.08	6.15	6.00±0.22	-0.15
W86 Hζ3	6.64	6.33±0.12	-0.31	6.73	6.46±0.07	-0.27
W86 Hε3	6.99	6.85±0.11	-0.14	7.03	6.90±0.06	-0.13
T90 Hα	4.24	4.10±0.09	-0.14	4.26	4.06±0.09	-0.20
T90 Hβ	4.56	4.18±0.12	-0.38	4.60	4.18±0.11	-0.42
T90 Hγ2	0.28	0.64±0.13	0.36	0.22	0.60±0.08	0.38
T142 Hα	3.63	3.78±0.02	0.15	3.63	3.78±0.03	0.15
T142 Hβ	3.43	3.83±0.20	0.40	3.43	3.96±0.06	0.53
T142 Hγ2	0.83	0.71±0.08	-0.08	0.83	0.76±0.03	-0.07
T178 Hα	3.87	4.08±0.04	0.21	3.77	4.05±0.04	0.28
T178 Hβ	4.08	4.25±0.03	0.17	4.00	4.24±0.03	0.24
T178 Hγ2	-0.05	0.35±0.06	0.40	-0.13	0.31±0.09	0.44
W182 Hδ1	7.43	7.22±0.03	-0.21	7.38	7.19±0.04	-0.19
W182 Hζ2	7.45	7.25±0.02	-0.20	7.27	7.26±0.01	-0.01
W182 Hη2	7.10	6.98±0.01	-0.12	7.19	6.97±0.01	-0.22
W182 Hζ3	7.07	6.97±0.03	-0.10	7.05	6.95±0.02	-0.10
W182 Hε3	7.52	7.39±0.02	-0.13	7.38	7.39±0.02	0.01

Tab. 3.10 Observed and average calculated chemical shifts of the assigned residues, which were allowed to move during the calculation.

3.7.3 Comparison of NMR and crystal structures

The present chapter will discuss the subtle differences between our NMR structures and the crystal structures of the all-trans ground state and two intermediates of the photocycle. A comparison of X-ray, electron microscopy and average NMR structures of the all-trans, 15-anti ground state are shown in Fig. 3.12 and Tab. 3.11.

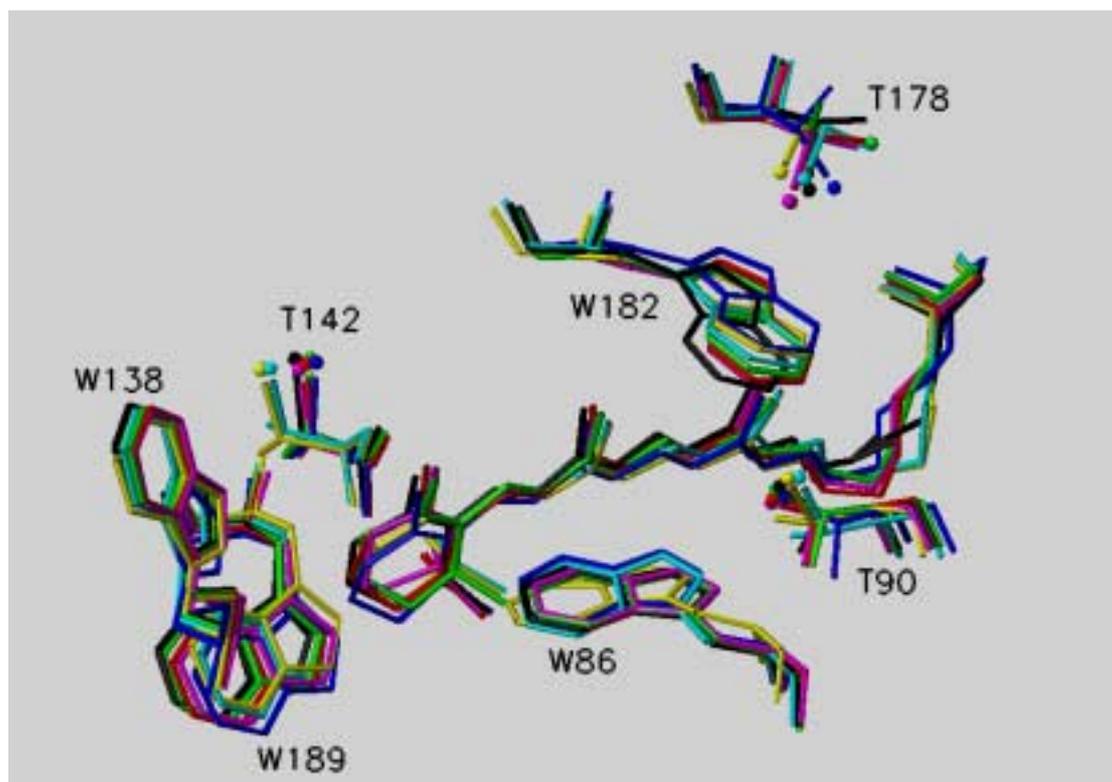


Fig. 3.12 Comparison of the average *all-trans* NMR structure (black) and the coordinates 1c3w (magenta), 1qhj (red), 1brr_B (green), 1bm1 (cyan), 2at9 (blue) and 2brd (yellow). The pseudo atoms representing the γ -methyl protons of the threonine side-chain are shown as spheres to allow distinguishing C γ and O γ .

<i>rms of the all-trans family relative to other structures</i>	<i>Ic3w</i>	<i>Iqhj</i>	<i>Ibrr</i>	<i>Ibm1</i>	<i>2at9</i>	<i>2brd</i>	
bb (sub)	0.33	0.27	0.17	0.41	0.59	0.51	
Ret	0.54	0.46	0.49	0.75	0.55	0.78	
	(0.36)	(0.29)	(0.27)	(0.33)	(0.33)	(0.33)	
Lys 216	1.02	1.08	0.96	0.99	1.03	1.09	
	(0.69)	(0.67)	(0.65)	(0.59)	(0.54)	(0.58)	
Trp 86	0.39	0.36	0.35	0.37	0.51	0.83	
Trp 138	0.33	0.37	0.27	0.29	0.69	0.38	
Trp 182	1.16	1.03	0.80	0.83	1.43	1.27	
Trp 189	0.62	0.44	0.29	0.82	0.92	0.84	
<i>pairwise rms</i>							
retinal ^{bb}	<i>av. nmr</i>	<i>Ic3w</i>	<i>Iqhj</i>	<i>Ibrr</i>	<i>Ibm1</i>	<i>2at9</i>	<i>2brd</i>
<i>av. nmr</i>		0.32	0.26	0.15	0.40	0.58	0.51
<i>Ic3w</i>	0.32		0.20	0.32	0.43	0.61	0.52
<i>Iqhj</i>	0.24	0.31		0.26	0.37	0.53	0.51
<i>Ibrr</i>	0.22	0.37	0.33		0.38	0.55	0.51
<i>Ibm1</i>	0.28	0.34	0.28	0.27		0.58	0.42
<i>2at9</i>	0.28	0.29	0.25	0.28	0.11		0.65
<i>2brd</i>	0.28	0.34	0.28	0.27	0.00	0.11	
<i>pairwise rms</i>							
K ₂₁₆ ^{W182}	<i>av. nmr</i>	<i>Ic3w</i>	<i>Iqhj</i>	<i>Ibrr</i>	<i>Ibm1</i>	<i>2at9</i>	<i>2brd</i>
<i>av. nmr</i>		1.15	1.01	0.77	0.80	1.41	1.25
<i>Ic3w</i>	0.80		0.17	0.47	0.60	0.63	0.59
<i>Iqhj</i>	0.88	0.23		0.32	0.46	0.64	0.57
<i>Ibrr</i>	0.71	0.34	0.39		0.27	0.73	0.62
<i>Ibm1</i>	0.88	0.90	0.86	0.77		0.69	0.48
<i>2at9</i>	0.81	0.82	0.80	0.57	0.98		0.48
<i>2brd</i>	0.90	0.99	0.94	0.88	0.20	1.06	

Tab. 3.11 Structure comparison of the all-trans family of structures and crystal structures. The first part of the table contains the average rms of selected coordinates (the error is very similar to that in Tab. 3.9 and is omitted). The 12 structures of the all-trans family have been fit to the core residues of each crystal structure. The values in parentheses are obtained by after a pairwise fit of the Ret carbon and Lys 216 sidechain atoms to the corresponding atoms of the crystal structures. The lower part of the table provides pairwise rms values for selected atoms between all crystal structures and the average NMR structure. In the middle part of the table, the structures are overlaid on the core backbone atoms (rms in upper right triangle) and retinal C atoms (rms in lower left triangle). In the lower part, the structures are overlaid on the core backbone atoms. The upper right triangle is the rms of the Trp 182 heavy sidechain atoms, the lower left the rms of the Lys 216 sidechain.

The largest deviation between the NMR family of structures and the crystal structures is found for the side-chain of Trp 182. The orientation of Trp 182 is similar in both the all-trans and the 13-cis, 15-syn families of NMR structures. Differences of the all-trans, 15-anti NMR structure to the crystal structures are therefore not caused by a mixing of the all-trans and 13-cis, 15-syn conformations in the crystals. (The crystal structures are usually recorded on samples, which have been illuminated and then frozen before the diffraction data are recorded. It is assumed that most of the proteins

are in the light adapted ground state (all-trans), but it cannot be excluded that a substantial amount of 13-cis, 15-syn is present as well.)

The possibility of an error in the NMR distance constraint lists for Trp 182 has been ruled out (see 3.7.1). The different orientation of the Trp side-chain is not caused by the NOE constraints, but reflects the minimal energy position of this side-chain as determined by the AMBER force-field.

The chemical shifts of the methyl group of Thr 178 are strongly influenced by the orientation of the Trp 182 side-chain. The calculated values for Thr 178 CH₃γ are -0.37 ppm (1c3w), -0.38 ppm (1qhj), 1.36 ppm (1brr), -0.08 ppm (1bm1), -0.02 ppm (2at9) and -0.08 ppm (2brd) for the crystal structures. The calculated average value for the twelve all-trans structures is 0.31±0.09 with a minimum value of 0.09. The average value of the crystal structures (excluding 1brr, which has the Thr 178 side-chain in false orientation), -0.19 ppm, is much closer to the experimental value of -0.13 ppm. Therefore the orientation of Trp 182 in the crystal structure is probably closer to the true structure in micelles than that of the NMR family of structures. The observed chemical shift in the 13-cis, 15-syn is -0.05 ppm and thus closer to the average of the calculated values for the all-trans crystal structures as well. The Trp 182 side-chain of the 13-cis, 15-syn conformer is expected to be in approximately the same orientation as the Trp 182 side-chain in the crystal structures. The small observed upfield shift of +0.08 ppm of the methyl of Thr 178 might indicate a small displacement of the Trp 182 side-chain in the direction of the NMR average positions for 13-cis, 15-syn as compared to all-trans.

The reason why the minimum energy conformation of the Trp 182 side-chain of the AMBER force field is slightly different from the true orientation is probably a missing interaction between the imino proton and a water molecule. Electron density for a water molecule, which is hydrogen bonded to both W182 Hε1 and A215 O was found by Luecke et al. (Luecke, 1999b). This interaction probably stabilizes the orientation of the Trp 182 side-chain closer to the backbone of helix G. There is no experimental evidence, that a water molecule is present in the late M-state structure at the same position, because the Trp 182 side-chain moves further away from the backbone of helix G. The observed distance in the average NMR structures (5.45 Å average all-

trans and 5.44 Å average 13-cis, 15-syn) is about halfway in between those of the all-trans (4.43 Å) and 13-cis, 15-anti (M state 6.99 Å).

It is very interesting to note, that the backbone C α and CO atoms of Trp 182 in the average NMR structures are closer to the 13-cis, 15-anti structure than to the all-trans structure. A possible interpretation of this would be that the favorable interaction between the Trp 182 side-chain and the water molecule causes conformational stress, which is partially released by a small local backbone distortion of helix F. This interpretation is favored by the observation, that the local backbone rms between different crystal structures shows a little increase in the region Trp 182, Ser 182 and Ala 184 compared to the neighboring residues in helix F. The electron density for the Trp side-chain is well resolved and so all crystal structures have this side-chain in the same conformation. However, the slightly distorted backbone conformation is more difficult to handle and the refined structures of the different groups therefore show larger backbone rms for that region.

The position of the retinal of the 13-cis, 15-syn structure is in between that of the all-trans ground state and the late M intermediate. The movement of C20 toward the cytoplasmic side is 1.62 Å in 1c3w relative to 1c8s, 0.32 Å in 1c3w relative average all-trans and 0.87 Å in 1c3w relative to average 13-cis, 15-syn. The distance of the methyl group in the 13-cis, 15-syn ground state to both crystal structures is approximately the same (0.87 Å to 1c3w and 0.84 to 1c8s), the distance to the average all-trans NMR structure is 0.56 Å. The movement of the retinal in the M structure is translated into a movement of the backbone atoms of Lys 216. This is not the case for the average NMR structures of both all-trans and 13-cis, 15-syn.

Very recently a high resolution X-ray structure of an early intermediate of the photocycle was determined (Edman, 1999). Diffraction data at 2.1 Å resolution on a mixture of the all-trans ground state and the low temperature K-intermediate were collected. The structure of the K intermediate was determined based on electron density differences. The retinal backbone and the Schiff base linkage up to C ϵ of Lys 216 of the all-trans and 13-cis, 15-anti conformers were constraint to be planar. The K intermediate (13-cis, 15-anti) and the average 13-cis, 15-syn structure show a very high similarity in the coordinates of the retinal starting from C14. The distance

between the C14 and C20 atoms is 0.20 Å and 0.35 Å respectively, if the two structures are aligned by minimizing the rms of the helical core residue (Subramaniam, 1999) backbone atoms. The Schiff base Nitrogen atoms are separated by more than one Angstrom and the water molecule making a hydrogen bond to the Schiff-base proton (W402 in 1c3w) is removed in the K intermediate.

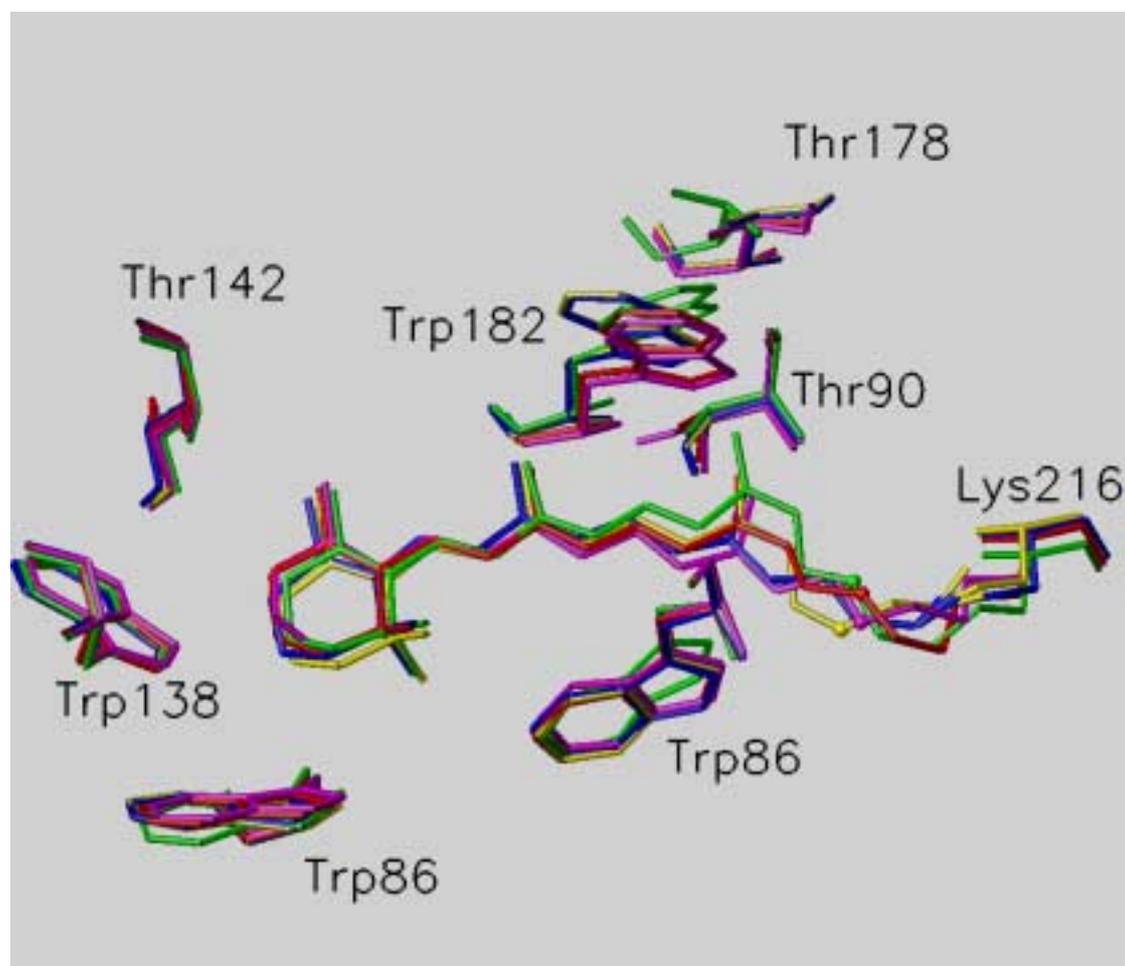


Fig. 3.13 Comparison of the average *all-trans* (blue) and 13-*cis*, 15-*syn* (yellow) NMR structures and the coordinates of the *all-trans* (magenta, 1c3w (Luecke, 1999b)), late M-intermediate (green, 1c8s (Luecke, 1999a)) and K-intermediate (red, 1qko (Edman, 1999))

3.7.4 Reflection on the BC and DE loops as observed in the crystal structures on the basis of NMR evidence

Additional assignments for two pairs of Trp and Thr residues have been obtained in chapter 3.4.3 (see also Tab. 3.6 of chapter 3.6.1), based on the analysis of distances and theoretical chemical shifts in the different crystal structures. These residues are located in the extracellular loop regions between helices B and C (Thr 67 and Trp 80) and helices D and E (Thr 121 and Trp 137) respectively, where the variation of the different crystal structures is larger than for the backbone atoms of the helices (Subramaniam, 1999). In the following, the loop structures of the different crystal structures are compared to NMR derived parameters.

Thr 67 is located in the loop between helices B and C (see Fig. 3.14). It is the first residue of a short β -strand formed by residues 67-71 and 74-78 (Luecke, 1999b). After allowing for ring-current shifts, the α protons of residues in helices tend to be shifted high field by approximately 0.30 ppm, whereas those in β -sheets tend to be 0.36 ppm down field shifted compared to the random coil shift (Ösapay and Case, 1994; Williamson, 1990; Wishart, 1992). This corresponds well to the observation, that the chemical shift of the α -proton of Thr 67 is shifted further down field than all other (assigned as well as unassigned) threonine resonances. The closest correspondence between the calculated H_{α} shift and its experimental value is observed for 1c3w (Luecke, 1999b).

The pattern (relative size) of the cross-peaks between Thr 67 and Trp 80 corresponds well to the distances in 1c3w (Luecke, 1999b), 1qhj (Belrhali, 1999) and 1brr (Essen, 1998) (see Tab. 3.6). These structures show relatively small deviations in the orientation of the Trp 80 and Thr 67 side-chain conformation as well as for the backbone in the BC loop region.

As judged from the few observed distances and chemical shifts, the region of the BC loop structure in solution is very similar to that observed in the crystal structures 1c3w (Luecke, 1999b), 1qhj (Belrhali, 1999) and 1brr (Essen, 1998). These structures show the lowest local backbone rms with respect to each other as compared to the other crystallographic structures.

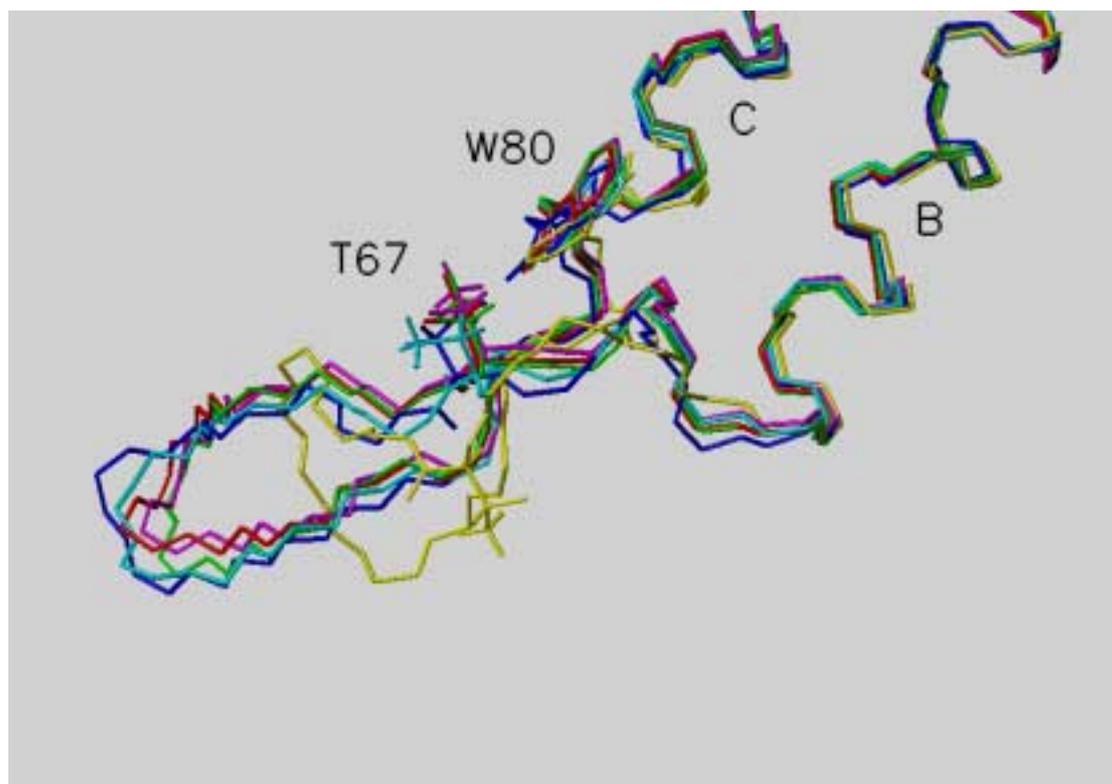


Fig. 3.14 Loop region between helices B and C in a superposition of bacteriorhodopsin coordinates 1c3w (magenta), 1qhj (red), 1brr_B (green), 1bm1 (cyan), 2at9 (blue) and 2brd (yellow). The structures have been superimposed on to 1c3w (used as a reference structure) minimizing the root mean squared deviation of the backbone atoms of the core of the transmembrane helices.

Thr 121 is located at the extracellular end of helix D (see Fig. 3.15). The backbone of all crystallographic structures is very similar in that region. However, the side-chain dihedral χ_1 of Thr 121 differs in the structures. Two groups with a similar rotamer are observed: 1c3w $\chi_1=49.9$, 1qhj $\chi_1= 51.9$, 2at9 $\chi_1=55.3$ and 1brr $\chi_1=-77.1$, 1bm1 $\chi_1=-78.3$, 2brd $\chi_1=-55.7$. The observed pattern of cross-peaks (see Tab. 3.6) identifies H β of Thr 121 being closest to the aromatic Trp protons of Trp 137. This holds true only in the first group of crystallographic structures. The NOE data therefore suggest that χ_1 of Thr 121 is close to 50.

In summary, the crystal structures of the extracellular loops BC and DE as observed in the highest resolution X-ray structures correspond very well to NMR distances and chemical shifts. There are no indications for differences in the three-dimensional protein structure caused by crystal contacts or protein/detergent interactions for these parts of the molecule.

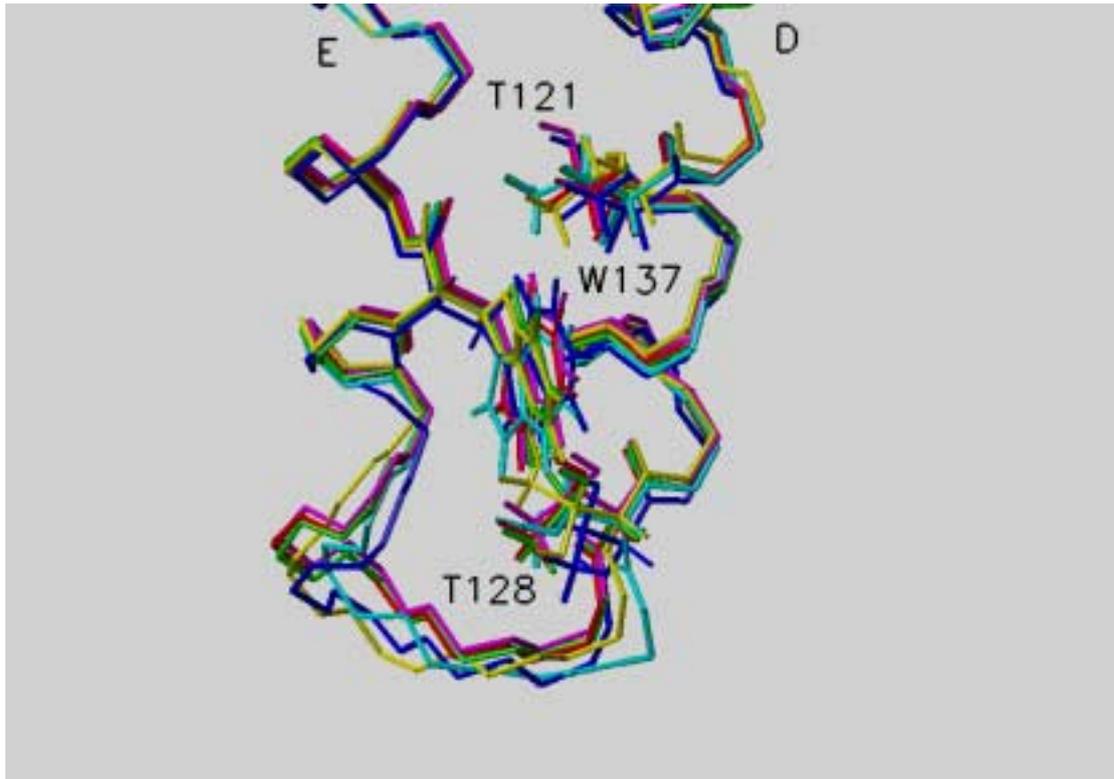


Fig. 3.15 Superposition of bacteriorhodopsin coordinates 1c3w (magenta), 1qhj (red), 1brr_B (green), 1bm1 (cyan), 2at9 (blue) and 2brd (yellow) for the extracellular part of helix D and E. The structures have been superimposed as in Fig. 3.14.

3.8 Discussion

The present chapter demonstrates that solution NMR can be successfully applied to perform structural studies on integral membrane proteins. Site specific assignments are possible and distance constraints can be obtained reliably. For the first time, a number of structurally important inter-helical distances for a membrane protein have been derived by solution NMR. The sequence specific assignments used for the structure calculation have to a large extent been made before the first high resolution X-ray structure (Pebay-Peyroula, 1997) of the protein has been published. The independent assignment of the retinal protons, which are subsequently used as a seed to identify nearby amino acids via a network of NOEs, was one of the keys to the success of the method. For the subsequent sequence specific assignment of the identified types of amino acids close to the retinal, a structural model is required. In the present case a model from electron crystallography, based on the experimentally observed density without further refinement (Henderson, 1990), was sufficient to identify Trp 86, 138, 182, 189 and Thr 142. The assignment of Thr 90 and 178 was

not unambiguous, because a third threonine (Thr 89) close in sequence and space brought in some bias. The crystal structures with improved resolution and NMR structure calculations with swapped assignments helped to sort out the correct assignments.

The labelling approach (selective ^1H on background ^2H) is used to the biggest advantage in cases, where unfavorable relaxation properties of the protein make heteronuclear experiments very difficult. For a *de novo* structure determination from NMR data alone, it may be used in addition to other recently developed concept for large proteins like TROSY and residual dipolar couplings. This is because it offers the possibility to determine long-range NOEs between amino acid side-chains, which are particularly important for NMR structure determination. The information is complementary to secondary structure constraints obtained from amide proton NOEs and orientational restraints from residual dipolar couplings in fully deuterated samples. In cases where structures are currently solved with samples fully ^{15}N ^{13}C labelled together with random fractional deuteration, the approach might be used to speed up the assignment and add additional distance constraints.

For structural investigations of solubilized membrane proteins, the above example outlines a general strategy. The labelling concept makes it possible to identify different amino acid types, which then in turn can be clustered into groups close in space. If it is possible to get sequence specific information of one member of this group, the neighbors can be identified from a structural model derived from other experimental data and/or molecular modeling. The starting model is then refined locally for the residues in the cluster. The crucial point is the sequence specific assignment of one residue of the cluster. In the present case the unique spin-system of the retinal could be used as starting point. It might be possible, that in proteins with only very few copies of one type of amino acid, the identification of different clusters makes it possible to assign them, however it should be kept in mind that the amount of information that can be obtained is very limited due to the disappearing of many signals under the noise level and the large overlap of the remaining ones caused by huge linewidth. A particular useful approach could be the combination of electron microscopy and NMR. The former method can produce three-dimensional maps at 7-9 Å resolution relatively quickly, but has more problems in achieving near atom

resolution, even if highly ordered two-dimensional crystals are available (Kühlbrandt and Gouaux, 1999). The method outlined in this chapter on the other hand has the potential to obtain atomic resolution for selected parts of the molecule, if low-resolution data are available.

The comparison of experimental chemical shifts and those calculated from high-resolution structures is shown to be a powerful and independent tool to evaluate the accuracy of structures. In the present case this approach is particularly fruitful because many of the residues including the retinal show large structural chemical shift. This leads to a very high sensitivity of these shifts to small changes in the atomic coordinates. The correlation of accuracy and chemical shift deviations is shown for the eleven crystal structures of all-trans bacteriorhodopsin available in the Brookhaven protein data bank. For the NMR structure determination, the overall chemical shift deviation was used to select a family of structures out of a larger number of simulated annealing runs. Since this is the first time that calculated chemical shifts are used in that way, the selected all-trans structures are compared to a family selected by minimum force field energy. It is demonstrated that the chemical shift selection results in both more accurate and more precise structures, while the force field criterion only increases the apparent precision and does not improve the accuracy.

Whether this approach is generally useful to select NMR structure out of a number of calculated structures remains to be shown. In an *ab initio* structure calculation, it could be applied to a subfamily of structures, from which members with very high energy representing models trapped in a local conformational minimum have been removed. The number of constraints is usually much higher and the additional chemical shift information might be unnecessary (Kuszewski, 1995). However, for regions with large structural shifts, a higher accuracy of the structures might result even in those cases and the additional computation time for a chemical shift selection is negligible compared to an approaches where proton chemical shift are used as constraints in every step of the simulated annealing protocol (Kuszewski, 1995; Ösapay and Case, 1991).

For the first time, a high-resolution structure of the 13-cis, 15-syn ground state of dark-adapted bacteriorhodopsin is presented. The structure is calculated under the assumption, that changes occur only locally in a region of 10 Å around the retinal, but

the overall arrangement of the seven helices is the same as for the all-trans ground state. The calculated structures for all-trans and 13-cis, 15-syn show no significant deviations in the side-chain conformations of the four tryptophan and three threonine residues, for which distance constraints have been determined. The largest changes upon cis-trans isomerization are observed for the C12-C15 moiety and the end of the side-chain of Lys 216. The Schiff-base nitrogen is brought closer in space to the Asp 85 and Asp 212 side-chains, which are part of the complex counterion (de Groot, 1989) formed by an interconnected branched network of six residues, four water molecules and 11 hydrogen bonds (Luecke, 1999b). The distances between Lys 216 N ζ and Asp85 O δ 2 and Asp 212 O δ 1 are reduced from 4.53 ± 0.49 Å and 3.93 ± 0.67 Å for the all-trans family of structure to 4.13 ± 0.51 Å and 3.44 ± 0.40 Å for 13-cis, 15-syn. In the high-resolution crystal structures, Asp85 O δ 2, Asp 212 O δ 1 and the protonated Schiff-base are all hydrogen bonded to same water-molecule. The stronger interaction of the protonated Schiff-base and its complex counterion for the 13-cis, 15-syn chromophore is in accord with solid state NMR measurement, which deduced the same from chemical shift differences of the Schiff-base nitrogen in the two dark adapted forms (de Groot, 1989).

The most significant difference between the NMR determined structures and the high-resolution crystal structures is a subtle change in the orientation of Trp 182. The orientation in the crystal structures is more accurate than that of the NMR structures as judged by chemical shift analysis. The difference is caused by a water molecule which is hydrogen bonded to both Trp 182 N ϵ 1 and Ala 215 O, which stabilizes the unusual side-chain conformation in the crystal structure (Kleywegt and Jones, 1998). This leads to a small local distortion of the α -helix backbone of residues 182-184 in the crystal structures, which is not observed in the NMR structures. In both NMR structures the stabilizing water molecule is not included in the structure calculation. The position of the C α atom of Trp 182 is closer to that observed in the X-ray structure of the late M state of the D96N mutant. In this intermediate of the photocycle, the Schiff-base has already transferred its proton to Asp 85 on the extracellular side and the switch of accessibility of the Schiff-base towards the cytoplasmic side has occurred, but the Schiff-base has not yet taken up a proton from the cytoplasmic side. Part of the structural changes of the protein in the M-state is the

movement of the Trp 182 side-chain towards the cytoplasmic side. The distance between its side-chain and Ala 215 is getting too big to be bridged by a water molecule.

Comparing the NMR 13-cis, 15-syn structure and the crystal structure of the K intermediate, in which the retinal is in a 13-cis, 15-anti conformation shows a surprising closeness of their C20 atoms. In the K intermediate, which arises within a few picoseconds following photoexcitation, the Schiff-base is still protonated and the structural changes in the protein are very small compared to the all-trans ground state. The steric interaction between the retinal C20 and Trp 182 was suggested to cause the structural changes observed at the cytoplasmic end of helix F later in the photocycle. From the comparison of the K structure and the 13-cis, 15-syn structure, such a mechanism seems unlikely, since in the later the arrangement of the Trp 182 side-chain and the retinal are very similar, yet it is a stable ground state structure.

Further improvements of the NMR structures should be achieved, if the chemical shifts are used directly as constraints during the refinement protocol (Kuszewski, 1995; Ösapay and Case, 1991). It is expected that with the inclusion of chemical shift restraints in the structure calculation, the number and accuracy of the distance constraints can be reduced, while structures of the same quality are achieved. This is particularly important for the study of intermediates of the photocycle, since the solubilized bacteriorhodopsin samples are not as stable under illumination as in the dark. If assignments are possible, it should also be possible to derive structural changes very accurately. This is because Trp 86 and 182 are two of the residues with large structural changes during the photocycle and it is expected that this leads to large chemical shift changes. The Trp 182 side-chain conformation has a strong influence on Thr 178 and Thr 90 and very small displacements of Trp 86 cause large changes in the shifts of its own protons as well as the retinal protons H10, H11 and H12.

The structural interpretation of the retinal chemical shift is going to be another big step to a better understanding of the protein. The ring-current contribution to the shifts of H10, H11 and H12 protons are larger than 2 ppm and thus larger than for any of the assigned proton resonances. A direct interpretation of these shifts has not been included in the present work, because in addition to the ring-currents, the changes in the π -electron system of the polyene chain have an influence on the proton chemical

shifts. However, the charge-distribution has a much stronger influence on carbon chemical shifts. Quantum chemical analysis of the chemical shifts of protons and carbons should therefore allow separating the contributions of local charge density on the proton shifts. Apart from gaining additional structural information, this opens as well the possibility to obtain accurate data on changes in the electron system of the retinal during the photocycle.

3.9 References

- Barsukov, I.L., Abdulaeva, G.V., Arseniev, A.S. and Bystrov, V.F. (1990) Sequence-specific ¹H-NMR assignment and conformation of proteolytic fragment 163-231 of bacterioopsin. *Eur J Biochem*, **192**, 321-7.
- Barsukov, I.L., Nolde, D.E., Lomize, A.L. and Arseniev, A.S. (1992) Three-dimensional structure of proteolytic fragment 163-231 of bacterioopsin determined from nuclear magnetic resonance data in solution. *Eur J Biochem*, **206**, 665-72.
- Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J.P., Landau, E.M. and Pebay-Peyroula, E. (1999) Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 Å resolution. *Structure*, **7**, 909-917.
- Bevington, P.R. and Robinson, D.K. (1992) *Data reduction and error analysis for the physical sciences*. Mc Graw Hill, Boston.
- Bodenhausen, G., Wagner, G., Rance, M., Sorensen, O.W., Wüthrich, K. and Ernst, R.R. (1984) *J. Magn. Reson.*, **59**, 542.
- Borgias, B.A., Gochin, M., Kerwood, D.J. and James, T.L. (1990) Relaxation matrix analysis of 2D NMR data. *Prog. Nucl. Magn. Reson. Spectrosc.*, **22**, 83-100.
- Bronstein, I.N., Semendjajew, K.A., Musiol, G. and Mühlig, H. (1997) *Taschenbuch der Mathematik*. Harri Deutsch, Frankfurt.
- Brünger, A.T. (1992) *X-PLOR Version 3.1 A system for X-ray crystallography and NMR*. Yale University Press, NewHaven.
- Case, D.A. (1995) Calibration of ring-current effects in proteins and nucleic acids. *J. Biomol. NMR*, **6**, 341-346.
- Cieplak, P., Cornell, W.D., Bayly, C.L. and Kollmann, P.A. (1995) *J. Comput. Chem.*, **16**, 1357-1377.
- Cornell, W.D., Cieplak, P., Bayly, C.L., Gould, I.R., Merz, K.M., Feguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollmann, P.A. (1995) A second generation force field for the simulation of proteins and nucleic acids. *J. Am. Chem. Soc.*, **117**, 5179-5197.
- Creuzet, F., McDermott, A., Gebhard, R., van der Hoef, K., Spijker-Assink, M.B., Herzfeld, J., Lugtenburg, J., Levitt, M.H. and Griffin, R.G. (1991) Determination of membrane protein structure by rotational resonance NMR: bacteriorhodopsin. *Science*, **251**, 783-6.
- de Groot, H.J., Harbison, G.S., Herzfeld, J. and Griffin, R.G. (1989) Nuclear magnetic resonance study of the Schiff base in bacteriorhodopsin: counterion effects on the ¹⁵N shift anisotropy. *Biochemistry*, **28**, 3346-53.
- de Groot, H.J., Smith, S.O., Courtin, J., van den Berg, E., Winkel, C., Lugtenburg, J., Griffin, R.G. and Herzfeld, J. (1990) Solid-state ¹³C and ¹⁵N NMR study of the low pH forms of bacteriorhodopsin. *Biochemistry*, **29**, 6873-83.
- Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R. and Landau, E.M. (1999) High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature*, **401**, 822-826.
- Efremov, R.G., Nolde, D.E., Vergoten, G. and Arseniev, A.S. (1999) A solvent model for simulations of peptides in bilayers. II. membrane- spanning alpha-helices. *Biophys J*, **76**, 2460-71.
- Essen, L., Siegert, R., Lehmann, W.D. and Oesterhelt, D. (1998) Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex. *Proc Natl Acad Sci U S A*, **95**, 11673-8.

- Farrar, M.R., Lakshmi, K.V., Smith, S.O., Brown, R.S., Raap, J., Lugtenburg, J., Griffin, R.G. and Herzfeld, J. (1993) Solid state NMR study of [epsilon-13C]Lys-bacteriorhodopsin: Schiff base photoisomerization. *Biophys J*, **65**, 310-5.
- Fletcher, C.M., Jones, D.N.M., Diamond, R. and Neuhaus, D. (1996) Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures. *J. Biomol. NMR*, **8**, 292-310.
- Glaubitx, C., Burnett, I.J., Gröbner, G., Mason, A.J. and Watts, A. (1999) Deuterium-MAS NMR spectroscopy on oriented membrane proteins: applications to photointermediates of bacteriorhodopsin. *J. Am. Chem. Soc.*, **121**, 5787-5794.
- Grabchuk, I.A., Orekhov, V. and Arseniev, A.S. (1996) 1H-15N backbone resonance assignments of bacteriorhodopsin. *Pharm Acta Helv*, **71**, 97-102.
- Griffin, R.G. (1998) Dipolar recoupling in MAS spectra of biological solids. *Nat. Struct. Biol. NMR suppl.*, **5**, 508-512.
- Grigorieff, N., Ceska, T.A., Downing, K.H., Baldwin, J.M. and Henderson, R. (1996) Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J Mol Biol*, **259**, 393-421.
- Grisshammer, R. and Tate, C.G. (1995) Overexpression of integral membrane proteins for structural studies. *Q. Rev. Biophys.*, **28**, 315-422.
- Haigh, C.W. and Mallion, R.B. (1980) *Prog. NMR Spectrosc.*, **13**, 303-344.
- Harbison, G.S., Herzfeld, J. and Griffin, R.G. (1983) Solid-state nitrogen-15 nuclear magnetic resonance study of the Schiff base in bacteriorhodopsin. *Biochemistry*, **22**, 1-4.
- Harbison, G.S., Smith, S.O., Pardo, J.A., Courtin, J.M., Lugtenburg, J., Herzfeld, J., Mathies, R.A. and Griffin, R.G. (1985) Solid-state 13C NMR detection of a perturbed 6-s-trans chromophore in bacteriorhodopsin. *Biochemistry*, **24**, 6955-62.
- Harbison, G.S., Smith, S.O., Pardo, J.A., Mulder, P.P., Lugtenburg, J., Herzfeld, J., Mathies, R. and Griffin, R.G. (1984a) Solid-state 13C NMR studies of retinal in bacteriorhodopsin. *Biochemistry*, **23**, 2662-7.
- Harbison, G.S., Smith, S.O., Pardo, J.A., Winkel, C., Lugtenburg, J., Herzfeld, J., Mathies, R. and Griffin, R.G. (1984b) Dark-adapted bacteriorhodopsin contains 13-cis, 15-syn and all-trans, 15-anti retinal Schiff bases. *Proc Natl Acad Sci U S A*, **81**, 1706-9.
- Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol*, **213**, 899-929.
- Henry, G.D. and Sykes, B.D. (1994) Methods to study membrane protein structure in solution, in *Methods in Enzymology* [T. L. James and N. J. Oppenheimer (eds.)], Academic Press, San Diego, Vol. 239, pp. 515-535.
- Hoch, J.C. and Stern, A.S. (1996) *NMR data processing*. Wiley, New York.
- Hu, J., Griffin, R.G. and Herzfeld, J. (1994) Synergy in the spectral tuning of retinal pigments: complete accounting of the opsin shift in bacteriorhodopsin. *Proc Natl Acad Sci U S A*, **91**, 8880-4.
- Hu, J.G., Sun, B.Q., Bizounok, M., Hatcher, M.E., Lansing, J.C., Raap, J., Verdegem, P.J., Lugtenburg, J., Griffin, R.G. and Herzfeld, J. (1998) Early and late M intermediates in the bacteriorhodopsin photocycle: a solid-state NMR study. *Biochemistry*, **37**, 8088-96.
- Hu, J.G., Sun, B.Q., Petkova, A.T., Griffin, R.G. and Herzfeld, J. (1997) The pre-discharge chromophore in bacteriorhodopsin: a 15N solid-state NMR study of the L photointermediate. *Biochemistry*, **36**, 9316-22.
- Jardetzky, O. and Roberts, G.C.K. (1981) *NMR in molecular biology*. Academic Press, New York.
- Keepers, J.W. and James, T.L. (1984) A theoretical study of distance determinations from NMR. Two-dimensional Overhauser effect spectra. *J. Magn. Reson.*, **57**, 404-426.
- Kimura, Y., Vassilyev, D.G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoaka, K., Murata, K., Hirai, T. and Fujiyoshi, Y. (1997) Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature*, **389**, 206-211.
- Kleywegt, G.J. and Jones, T.A. (1998) Databases in protein crystallography. *Acta Cryst.*, **D54**, 1119-1131.
- Konradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 51-55.
- Kühlbrandt, W. and Gouaux, E. (1999) Membrane proteins. Editorial overview. *Curr. Opin. Struct. Biol.*, **9**, 445-447.

- Kuszewski, J., Gronenborn, A.M. and Clore, G.M. (1995) The impact of direct refinement against proton chemical shifts on protein structure determination by NMR. *J. Magn. Reson. Ser. B*, **107**, 293-297.
- Lakshmi, K.V., Farrar, M.R., Raap, J., Lugtenburg, J., Griffin, R.G. and Herzfeld, J. (1994) Solid state ¹³C and ¹⁵N NMR investigations of the N intermediate of bacteriorhodopsin. *Biochemistry*, **33**, 8853-7.
- Landau, E.M. and Rosenbusch, J.P. (1996) Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci. USA*, **93**, 14532-14535.
- Landy, S.B. and Rao, B.D.N. (1989) Influence of molecular geometry on uncertainty in distances determined from NOE. *J. Magn. Reson.*, **83**, 29-43.
- LeMaster, D.M. (1994) Isotope labeling in solution protein assignment and structural analysis. *Prog. Nucl. Magn. Reson. Spectr.*, **26**, 371-419.
- Lomize, A.L., Pervushin, K.V. and Arseniev, A.S. (1992) Spatial structure of (34-65) bacterioopsin polypeptide in SDS micelles determined from nuclear magnetic resonance data. *J Biomol NMR*, **2**, 361-72.
- Losonczi, J.A. and Prestegard, J.H. (1998) Nuclear magnetic resonance characterization of the myristoylated, N- terminal fragment of ADP-ribosylation factor 1 in a magnetically oriented membrane array. *Biochemistry*, **37**, 706-16.
- Luecke, H., Richter, H.T. and Lanyi, J.K. (1998) Proton transfer pathways in bacteriorhodopsin at 2.3 Ångstrom resolution. *Science*, **227**, 1676-1681.
- Luecke, H., Schobert, B., Richter, H.T., Cartailler, J.P. and Lanyi, J.K. (1999a) Structural changes in bacteriorhodopsin during ion transport at 2 Ångstrom resolution. *Science*, **286**, 255-260.
- Luecke, H., Schobert, B., Richter, H.T., Cartailler, J.P. and Lanyi, J.K. (1999b) Structure of Bacteriorhodopsin at 1.55 Å Resolution. *J Mol Biol*, **291**, 899-911.
- Madrid, M., Mace, J.E. and Jardetzky, O. (1989) Consequences of magnetization transfer on the determination of solution structures of proteins. *J. Magn. Reson.*, **83**, 267-278.
- McConnell, H.M. (1957) *J. Chem. Phys.*, **27**, 226-229.
- McDermott, A.E., Creuzet, F., Gebhard, R., van der Hoef, K., Levitt, M.H., Herzfeld, J., Lugtenburg, J. and Griffin, R.G. (1994) Determination of internuclear distances and the orientation of functional groups by solid-state NMR: rotational resonance study of the conformation of retinal in bacteriorhodopsin. *Biochemistry*, **33**, 6129-36.
- Michel, H. (1991) General and practical aspects of membrane protein crystallization, in *Crystallization of membrane proteins* [H. Michel (ed.) CRC Press, Boca Raton.
- Mitsuoka, K., Hirai, T., Murata, K., Miyazawa, A., Kidera, A., Kimura, Y. and Fujiyoshi, Y. (1999) The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: implication of the charge distribution. *J Mol Biol*, **286**, 861-82.
- Moltke, S., Nevzorov, A.A., Sakai, N., Wallat, I., Job, C., Nakanishi, K., Heyn, M.P. and Brown, M.F. (1998) Chromophore orientation in bacteriorhodopsin determined from the angular dependence of deuterium nuclear magnetic resonance spectra of oriented purple membranes. *Biochemistry*, **37**, 11821-11835.
- Neuhaus, D. and Williamson, M. (1989) *The nuclear Overhauser effect in structural and conformational analysis*. VCH, New York.
- Nilges, M., Habazettl, J., Brünger, A.T. and Holak, T.A. (1991) Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J. Mol. Biol.*, **219**, 499-510.
- Nolde, D.E., Arseniev, A.S., Vergoten, G. and Efremov, R.G. (1997) Atomic solvation parameters for proteins in a membrane environment. Application to transmembrane alpha-helices. *J Biomol Struct Dyn*, **15**, 1-18.
- Opella, S.J. (1994) Solid-state NMR structural studies of proteins. *Annu Rev Phys Chem*, **45**, 659-83.
- Opella, S.J. (1997) NMR and membrane proteins. *Nat Struct Biol*, **4 Suppl**, 845-8.
- Opella, S.J., Kim, Y. and McDonnell, P. (1994) Experimental nuclear magnetic resonance studies of membrane proteins, in *Methods in Enzymologie* [T. L. James and N. J. Oppenheimer (eds.)], Academic Press, San Diego, Vol. 239, pp. 536-560.
- Opella, S.J., Marassi, F.M., Gesell, J.J., Valente, A.P., Kim, Y., Oblatt-Montal, M. and Montal, M. (1999) Structures of the M2 channel-lining segments from nicotinic acetylcholine and NMDA receptors by NMR spectroscopy. *Nat Struct Biol*, **6**, 374-9.
- Orekhov, V., Abdulaeva, G.V., Musina, L. and Arseniev, A.S. (1992) ¹H-¹⁵N-NMR studies of bacteriorhodopsin Halobacterium halobium. Conformational dynamics of the four-helical bundle. *Eur J Biochem*, **210**, 223-9.

- Orekhov, V., Pervushin, K.V. and Arseniev, A.S. (1994) Backbone dynamics of (1-71)bacterioopsin studied by two-dimensional ^1H - ^{15}N NMR spectroscopy. *Eur J Biochem*, **219**, 887-96.
- Ösapay, K. and Case, D.A. (1991) A new analysis of proton chemical shifts in proteins. *J. Am. Chem. Soc.*, **113**, 9436-9444.
- Ösapay, K. and Case, D.A. (1994) Analysis of proton chemical shifts in regular secondary structure of proteins. *J. Biomol. NMR*, **4**, 215-230.
- Pashkov, V.S., Balashova, T.A., Zhemaeva, L.V., Sikilinda, N.N., Kutuzov, M.A., Abdulaev, N.G. and Arseniev, A.S. (1996) Conformation of surface exposed N-terminus part of bacteriorhodopsin studied by transferred NOE technique. *FEBS Lett*, **381**, 119-22.
- Patel, D.J. (1969) 220 MHz proton nuclear magnetic resonance spectra of retinals. *Nature*, **221**, 825-828.
- Pattaroni, C.a.L., J (1981) A study of N-butyl-(all-trans-retinylidene)amine and its protonated species by ^1H - and ^{13}C -NMR spectroscopy. *Helv. Chim. Acta.*, **64**, 1969-1984.
- Patzelt, H., Ulrich, A.S., Egbringhoff, H., Düx, P., Ashurst, J., Simon, B., Oschkinat, H. and Oesterheld, D. (1997) Towards structural investigations on isotope labelled native bacteriorhodopsin in detergent micelles by solution-state NMR spectroscopy. *J. Biomol. NMR*, **10**, 95-106.
- Pebay-Peyroula, E., Rummel, G., Rosenbusch, J.P. and Landau, E.M. (1997) X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science*, **277**, 1676-81.
- Pervushin, K.V. and Arseniev, A.S. (1992) Three-dimensional structure of (1-36)bacterioopsin in methanol- chloroform mixture and SDS micelles determined by 2D ^1H -NMR spectroscopy. *FEBS Lett*, **308**, 190-6.
- Pervushin, K.V., Orekhov, V., Popov, A.I., Musina, L. and Arseniev, A.S. (1994) Three-dimensional structure of (1-71)bacterioopsin solubilized in methanol/chloroform and SDS micelles determined by ^{15}N - ^1H heteronuclear NMR spectroscopy. *Eur J Biochem*, **219**, 571-83.
- Post, C.B., Meadows, R.P. and Gorenstein, D.G. (1990) On the evaluation of interproton distances for three-dimensional structure determination by NMR using a relaxation rate matrix analysis. *J. Am. Chem. Soc.*, **112**, 6796-6803.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) *Numerical recipes in C, The art of scientific computing*. Cambridge University Press, Cambridge.
- Rowan III, R., Warshel, A., Sykes, B. D. and Karplus, M (1972) Conformation of retinal isomers. *Biochemistry*, **13**, 970-980.
- Rowan III, R. and Sykes, B.D. (1975) A ^1H nuclear magnetic resonance determination of the conformations of the polyene chain portions of 9-cis- and 13-cis-retinal in solution. *J. Am. Chem. Soc.*, **97**, 1023-1027.
- Sanders, C.R., Hare, B.J., Howard, K.P. and Prestegard, J.H. (1994) Magnetically-oriented phospholipid micelles as a tool for the study of membrane associated molecules. *Prog. NMR Spectrosc.*, **26**, 421-444.
- Sanders, C.R. and Landis, G.S. (1995) Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry*, **34**, 4030-4040.
- Sass, H.J., Berendzen, J., Neff, D., Gessenich, R., Ormos, P. and Bueldt, G. (20-OCT-99) The M intermediate structure of bacteriorhodopsin at 2.2 Å resolution. *to be published*.
- Sato, H., Takeda, K., Tani, K., Hino, T., Okada, T., Nakasako, M., Kamiya, N. and Kouyama, T. (1999) Specific lipid-protein interactions in a novel honeycomb lattice structure of bacteriorhodopsin. *Acta Crystallogr D Biol Crystallogr*, **55**, 1251-6.
- Schmid, M.W., Baldrige, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S.J., Windus, T.L., Dupuis, M. and Montgomery, J.A. (1993) *J. Comput. Chem.*, **14**, 1347-1363.
- Schmieder, P. (1991) *Entwicklung und Anwendung neuer heteronuklearer NMR-spektroskopischer Methoden mit Protonendetektion*. Promotionsschrift TU München.
- Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Groger, A., Schwarz, C., Schuldiner, S. and Kessler, H. (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur J Biochem*, **254**, 610-9.
- Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, L. (1991) A relational database for sequence-specific protein NMR data. *J. Biomol. NMR*, **1**, 217-236.
- Seigneuret, M. and Kainosho, M. (1993) Localization of methionine residues in bacteriorhodopsin by carbonyl ^{13}C -NMR with sequence-specific assignments. *FEBS Lett*, **327**, 7-12.

- Seigneuret, M., Neumann, J.M., Levy, D. and Rigaud, J.L. (1991a) High-resolution ^{13}C NMR study of the topography and dynamics of methionine residues in detergent-solubilized bacteriorhodopsin. *Biochemistry*, **30**, 3885-92.
- Seigneuret, M., Neumann, J.M. and Rigaud, J.L. (1991b) Detergent delipidation and solubilization strategies for high-resolution NMR of the membrane protein bacteriorhodopsin. *J Biol Chem*, **266**, 10066-9.
- Smith, S.O., Aschheim, K. and Groesbeek, M. (1996) Magic angle spinning NMR spectroscopy of membrane proteins. *Quart. Rev. Biophys.*, **29**, 395-449.
- Smith, S.O., Courtin, J., van den Berg, E., Winkel, C., Lugtenburg, J., Herzfeld, J. and Griffin, R.G. (1989a) Solid-state ^{13}C NMR of the retinal chromophore in photointermediates of bacteriorhodopsin: characterization of two forms of M. *Biochemistry*, **28**, 237-43.
- Smith, S.O., de Groot, H.J., Gebhard, R., Courtin, J.M., Lugtenburg, J., Herzfeld, J. and Griffin, R.G. (1989b) Structure and protein environment of the retinal chromophore in light- and dark-adapted bacteriorhodopsin studied by solid-state NMR. *Biochemistry*, **28**, 8897-904.
- Sobol, A.G., Arseniev, A.S., Abdulaeva, G.V., Musina, L. and Bystrov, V.F. (1992) Sequence-specific resonance assignment and secondary structure of (1-71) bacterioopsin. *J Biomol NMR*, **2**, 161-71.
- Stejskal, E.O.M., J. D. (1994) *High Resolution NMR in the Solid State*. Oxford University Press, New York.
- Subramaniam, S. (1999) The structure of bacteriorhodopsin: an emerging consensus. *Curr Opin Struct Biol*, **9**, 462-8.
- Thompson, L.K., McDermott, A.E., Raap, J., van der Wielen, C.M., Lugtenburg, J., Herzfeld, J. and Griffin, R.G. (1992) Rotational resonance NMR study of the active site structure in bacteriorhodopsin: conformation of the Schiff base linkage. *Biochemistry*, **31**, 7931-8.
- Tycko, R. (1996) Prospects for resonance assignments in multidimensional solid-state NMR spectra of uniformly labeled proteins. *J. Biomol. NMR*, **8**, 239-251.
- Ulrich, A.S., Heyn, M.P. and Watts, A. (1992) Structure determination of the cyclohexene ring of retinal in bacteriorhodopsin by solid-state deuterium NMR. *Biochemistry*, **31**, 10390-9.
- Ulrich, A.S., Wallat, I., Heyn, M.P. and Watts, A. (1995) Re-orientation of retinal in the M-photointermediate of bacteriorhodopsin. *Nat. Struct. Biol.*, **2**, 190-192.
- Ulrich, A.S., Watts, A., Wallat, I. and Heyn, M.P. (1994) Distorted structure of the retinal chromophore in bacteriorhodopsin resolved by ^2H -NMR. *Biochemistry*, **33**, 5370-5.
- Williams, K.A., Farrow, N.A., Deber, C.M. and Kay, L.E. (1996) Structure and dynamics of bacteriophage λ major coat protein in MPG micelles by solution NMR. *Biochemistry*, **35**, 5145-57.
- Williamson, M.P. (1990) Secondary-structure dependent chemical shifts in proteins. *Biopolymers*, **29**, 1423-1431.
- Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, **5**, 67-81.
- Wishart, D.S., Richards, F.M. and Sykes, B.D. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, **31**, 1647-1651.
- Wishart, D.S. and Sykes, B.D. (1994) Chemical shifts as a tool for structure determination, in *Methods in Enzymology* [N. J. O. a. T. L. James (ed.) Academic Press, San Diego, Vol. 239, pp. 363-392.
- Wolfram, S. (1996) *MATHEMATICA*. Wolfram Research.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley-Interscience, New York.
- Yip, P. and Case, D.A. (1989) A new method for refinement of macromolecular structures based on nuclear Overhauser effect spectra. *J. Magn. Reson.*, **83**, 643-648.
- Zhao, D. and Jardetzky, O. (1994) An assessment of the precision and accuracy of protein structures by NMR. Dependence on distance errors. *J. Mol. Biol.*, **239**, 601-607.
- Zhou, F., Windemuth, A. and Schulten, K. (1993) Molecular dynamics study of the proton pump cycle of bacteriorhodopsin. *Biochem.*, **32**, 2291-2306.