Biochimica et Biophysica Acta 1466 (2000) 1–6

www.elsevier.com/locate/bba

Rapid report

Solution structure of the loops of bacteriorhodopsin closely resembles the crystal structure

Madan Katragadda [a], James L. Alderfer [b], Philip L. Yeagle [a,]*

[a] Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA
[b] Department of Biophysics, Roswell Park Cancer Institute, Buffalo, NY 14263, USA

Received 29 December 1999; received in revised form 27 January 2000; accepted 4 February 2000

Abstract

Bacteriorhodopsin is one of very few transmembrane proteins for which high resolution structures have been solved. The structure shows a bundle of seven helices connected by six turns. Some turns in proteins are stabilized by short range interactions and can behave as small domains. These observations suggest that peptides containing the sequence of the turns in a membrane protein such as bacteriorhodopsin may form stable turn structures in solution. To test this hypothesis, we determined the solution structure of three peptides each containing the sequence of one of the turns in bacteriorhodopsin. The solution structures of the peptides closely resemble the structures of the corresponding turns in the high resolution structures of the intact protein.

Keywords: Bacteriorhodopsin; NMR; Membrane protein; Structure

The paucity of high resolution structures for integral membrane proteins is a measure of the difficulty in obtaining structural information. In contrast to the family of soluble proteins, the structures of only a few transmembrane proteins have been solved by X-ray crystallography and nuclear magnetic resonance (NMR). This has inspired a search for alternate approaches by which some structural information can be obtained for integral membrane proteins.

One such approach is to determine the structures of discrete subdomains of the membrane protein rather than attempt the structure determination of the entire intact protein. Some classes of integral membrane proteins can be thought of as bundles of helices connected by turns. Turns and helices are normally stabilized by relatively short range interactions. The turns and the helices of some membrane proteins might therefore constitute such discrete subdomains.

Bacteriorhodopsin from *Halobacterium halobium* offers an opportunity to examine this hypothesis. Several crystal structures are available for this membrane protein. The structure of bacteriorhodopsin consists of seven transmembrane helices connected by turns. It has already been shown that most of the helices of bacteriorhodopsin show helical structures as individual peptides [1]. The turns of bacteriorhodopsin have not been examined. Of the six turns, three (CD, DE and FG) are defined in the crystal structures (the other three turns are poorly defined in the available data). We synthesized three peptides containing the amino acid sequences of each

---

* Corresponding author. Fax: +1-860-486-4331;
E-mail: yeagle@uconnvm.uconn.edu

PII: S0005-2736(00)00167-X

of these turns. The structures of these peptides in solution were determined by two dimensional homonuclear NMR.

The turns in the bacteriorhodopsin structure can be considered to be part of a helix-turn-helix motif. The sequences in this work were selected to correspond to the turns observed in the crystal structure, with some additional sequence on each side of the turn that in the crystal structures formed part of the helices that connect to each turn. A 23mer with the sequence LLDLALLVDADQGTILALVGADG (bacteriorhodopsin residues 90–112 (numbering from crystal structure [4])) was synthesized corresponding to the CD loop of bacteriorhodopsin. A 15mer with the sequence LVGALTKVYSYQFVW (bacteriorhodopsin residues 119–133) was synthesized corresponding to the DE loop of bacteriorhodopsin. A 24mer with the sequence PVVWLIGSEGAGIVPLNIETLLFM (bacteriorhodopsin residues 177–200) was synthesized corresponding to the FG loop of bacteriorhodopsin. These peptides were synthesized by solid phase synthesis in the Biotechnology Center at the University of Connecticut. These peptides are relatively hydrophobic and are not soluble in aqueous media for the length of the NMR experiments. DMSO effectively dissolved the peptides and these solutions were stable for the entire period of data collection.

All NMR spectra were recorded on a Bruker AMX-600 spectrometer at 30°C. Solutions were 2–3 mM in peptide. Standard pulse sequences and phase cycling were employed to record homonuclear ^{1}H NMR double quantum filtered COSY and NOESY (400 ms mixing time) spectra [2]. All spectra were accumulated in a phase sensitive manner using time-proportional phase incrementation for quadrature detection in F1. Chemical shifts were referenced to internal water in the DMSO.

The sequence-specific assignment of the ^{1}H NMR spectrum for the peptide was carried out using standard methods employing FELIX (MSI). Assigned NOE cross peaks were segmented using a statistical segmentation function and characterized as strong medium and weak corresponding to upper bounds distance range constraints of 2.7, 3.5 and 5.0 Å, respectively. Lower bounds between non-bonded atoms were set to the sum of their van der Waals radii (approximately 1.8 Å). Pseudoatom corrections

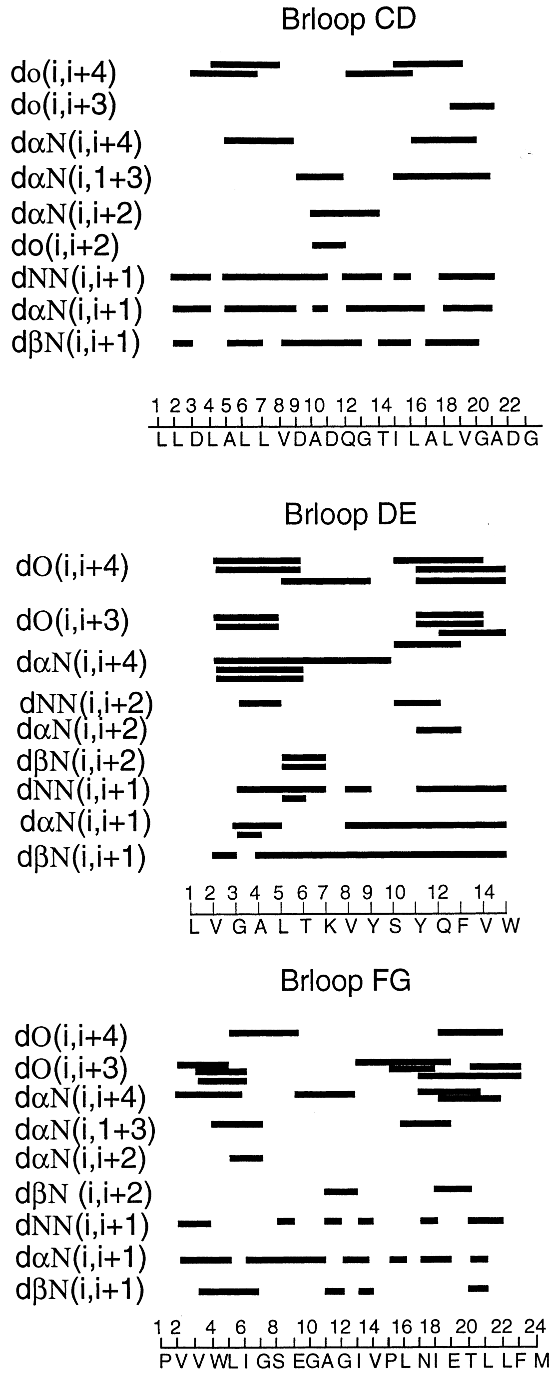


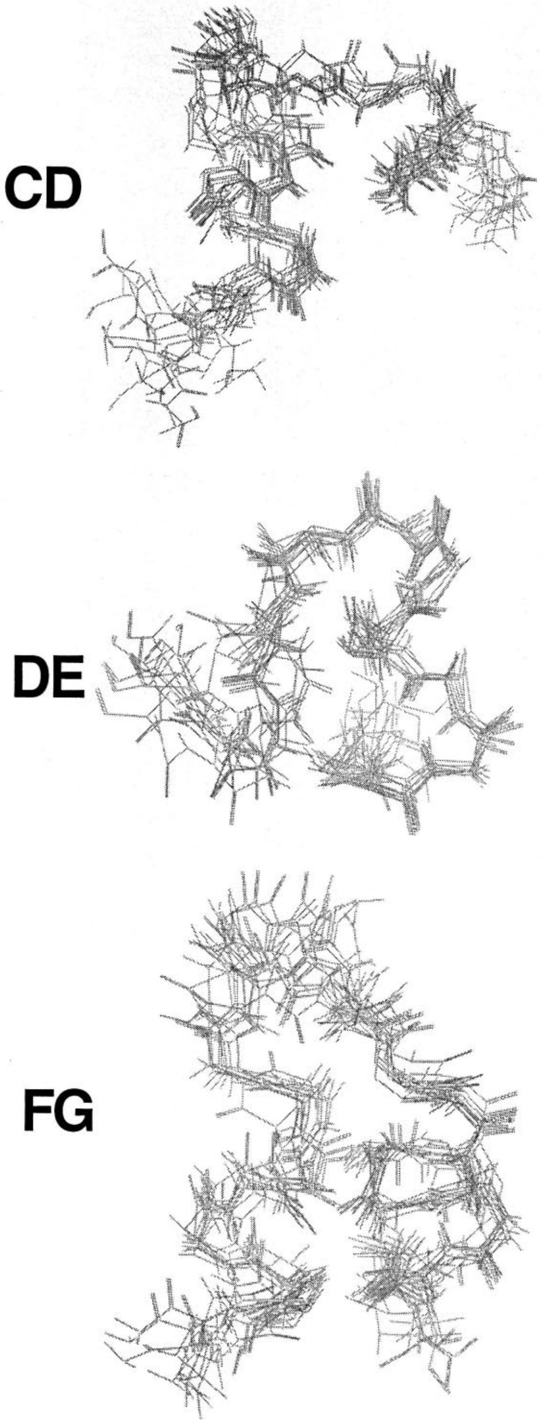
Fig. 1. Connectivity observed in the NOESY maps for each of the loop peptides. The sequences written are numbered internally within the peptide. The corresponding bacteriorhodopsin sequences are noted in the text. At the top of each panel, interactions other than αH–NH, βH–NH or NH–NH are also indicated ('do'). Clusters of i to i+3 and i to i+4 interactions are observed at the first third and last third of each peptide, consistent with the α-helix observed in the structures.

Table 1
Energy minimization details (kcal/mol)

Energy	CD loop	DE loop	FG loop
Bond stretching	22.327	18.485	13.326
Angle bending	135.880	233.599	138.662
Torsional	31.477	32.808	47.780
Improper torsional	3.264	4.701	5.580
1–4 van der Waals	79.575	80.901	63.423
van der Waals	−25.159	−1.437	−72.435
1–4 Electrostatic	338.077	191.937	338.995
Electrostatic	−610.193	−326.821	−580.496
H-Bond	−3.574	−0.507	−6.608
Fixed torsion angle	0.086	0.108	0.146
Fixed distance range	226.543	221.903	108.618
Total	198.303	455.677	56.992

were added to inter-proton distance restraints where necessary [3]. Structures were obtained using simulated annealing on the peptides with the distance constraints obtained from the NOESY data. Simulated annealing was done with the Kollman All Atom force field and Kollman charges, within SYBYL (Tripos). The molecule was heated to 1000 K for 1000 fs followed by cooling to 200 K during 1500 fs. Ten cycles were calculated. Statistics on structures were obtained from X-plor. These calculations were performed on a Silicon Graphics O2 workstation. Imaging and superposition of the resulting structures were performed on a Power Mac with MacImdad (Molecular Applications Group, Palo Alto, CA, USA) and on SYBYL using Fit Monomers. Energy minimization was performed in SYBYL with the Kollman All Atom force field and Kollman charges. SYBYL ProTable and ProCheck were used to analyze the structures.

The solution structures of these peptides were solved. For the CD loop, a total of 224 constraints were obtained: 72 intraresidue, 128 sequential and 44 long range. For the DE loop, a total of 172 constraints were obtained: 97 intraresidue, 54 sequential and 21 long range. For the FG loop, a total of 236 constraints were obtained: 103 intraresidue, 81 sequential and 52 long range. Fig. 1 shows the connectivity among these constraints for each peptide. Fig. 2 shows the superposition of a family of 10 structures obtained from simulated annealing for each peptide. The average structure was calculated and the average rmsd for the families from the average structure for each peptide was: CD 1.2; DE 0.80; FG 0.92. Table



→

Fig. 2. Families of 10 structures obtained for each of the peptides using simulated annealing as described in the text. Only the backbone atoms are shown.

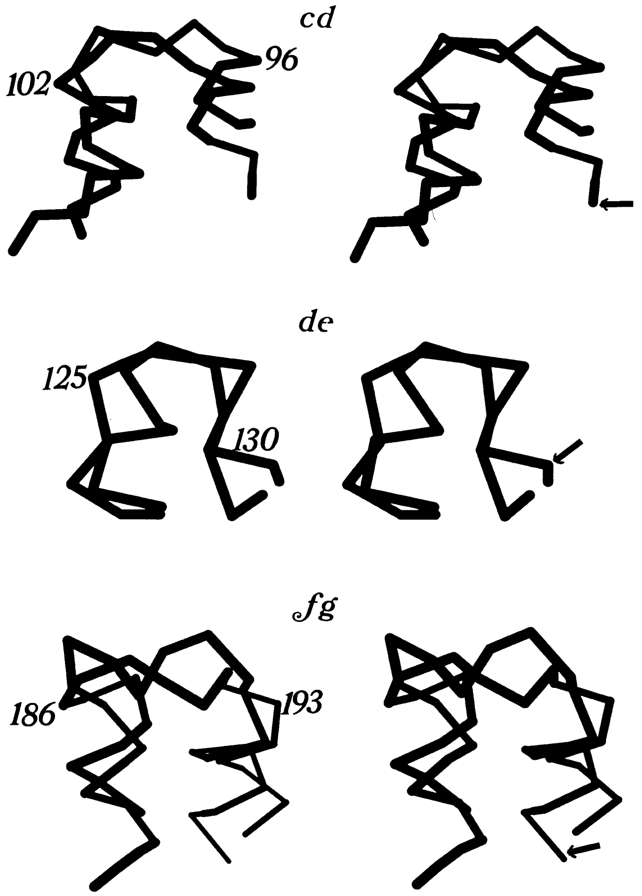


Fig. 3. Superposition of the structure for each loop peptide, CD, DE and FG, on the corresponding amino acid sequence in a recent crystal structure of bacteriorhodopsin (1c3w) [4]. Only the relevant loop region of the bacteriorhodopsin structure is represented for clarity. The arrow indicates the peptide whose structure is reported here. In each case, the peptide structure mimics the crystal structure except at the last two or three residues which are disordered in the solution structures as can be seen in Fig. 2. The amino-terminal is at the left and the carboxyl-terminal is at the right.

1 shows the energy analyses for these structures. ProCheck and ProTable were used to examine each peptide. Stereochemistry and configuration for each reside of each of the structures were normal, and out of all the ϕ–χ angles in all the peptide structures, only two were in an unacceptable region of the Ramachandran plot.

For each of the peptides representing the turns of bacteriorhodopsin, CD, DE and FG, a turn structure is observed in solution, reflecting the parent protein structure. The turn appears to be part of a helix-turn-helix motif. These turn structures are largely determined by short range distance constraints ($i \rightarrow i+1$, $i \rightarrow i+2$, $i \rightarrow i+3$ and $i \rightarrow i+4$), but there are a few long range constraints between the two sides of the turn as well. These latter constraints were helix–helix interactions. When structures were calculated in the absence of these latter constraints, the turn was still observed in the structure. Thus the inter-helical constraints did not artificially induce a turn in the structure calculations. However, the effect of including these constraints was to stabilize the relationship between the two helices during the calculations. These constraints were used in the structures reported here.

The loop structures from the peptides were superimposed on the corresponding amino acid sequence in a recent crystal structure of bacteriorhodopsin (1c3w) [4]. Residues 8–14 (126–132 of bR) of loop DE were superimposed on the crystal structure with an rmsd of two. The loop in this crystal structure is centered on residues 125–128. Thus the region superimposed includes almost the entire loop and much of the carboxyl-terminal helix of this structure. The region 120–126 was also superimposed with an rmsd of 2.5, corresponding to the helix on the amino-terminal side of this peptide. Residues at the very termini of the peptide were not used as those residues are disordered in the NMR structures of all the peptides. A previous analysis showed that an rmsd of 2.5 or less is observed when comparing peptides of differing length but encompassing a substantial structural component in common [5]. These data indicate good agreement between the crystal structure of bacteriorhodopsin and the DE loop peptide when compared segmentally. The rmsd goes up to about three if both helices and the loop are superimposed simultaneously. The relationship between the helices on both sides of the loop is possibly underdetermined by the available data, leading to a reduction in the quality of the fit. However, Fig. 3 shows that such a simultaneous superposition actually gives relatively good agreement with the crystal structure in that the size of the turn and general direction of the chain is the same in the peptide as in the crystal structure. Furthermore, the solution structure of the peptide is within the envelope of conformations represented in the available crystal structures, as shown in Fig. 4.

Residues 10–22 (186–198 of bR), or 13 residues, of the FG loop were superimposed on 1c3w, with an

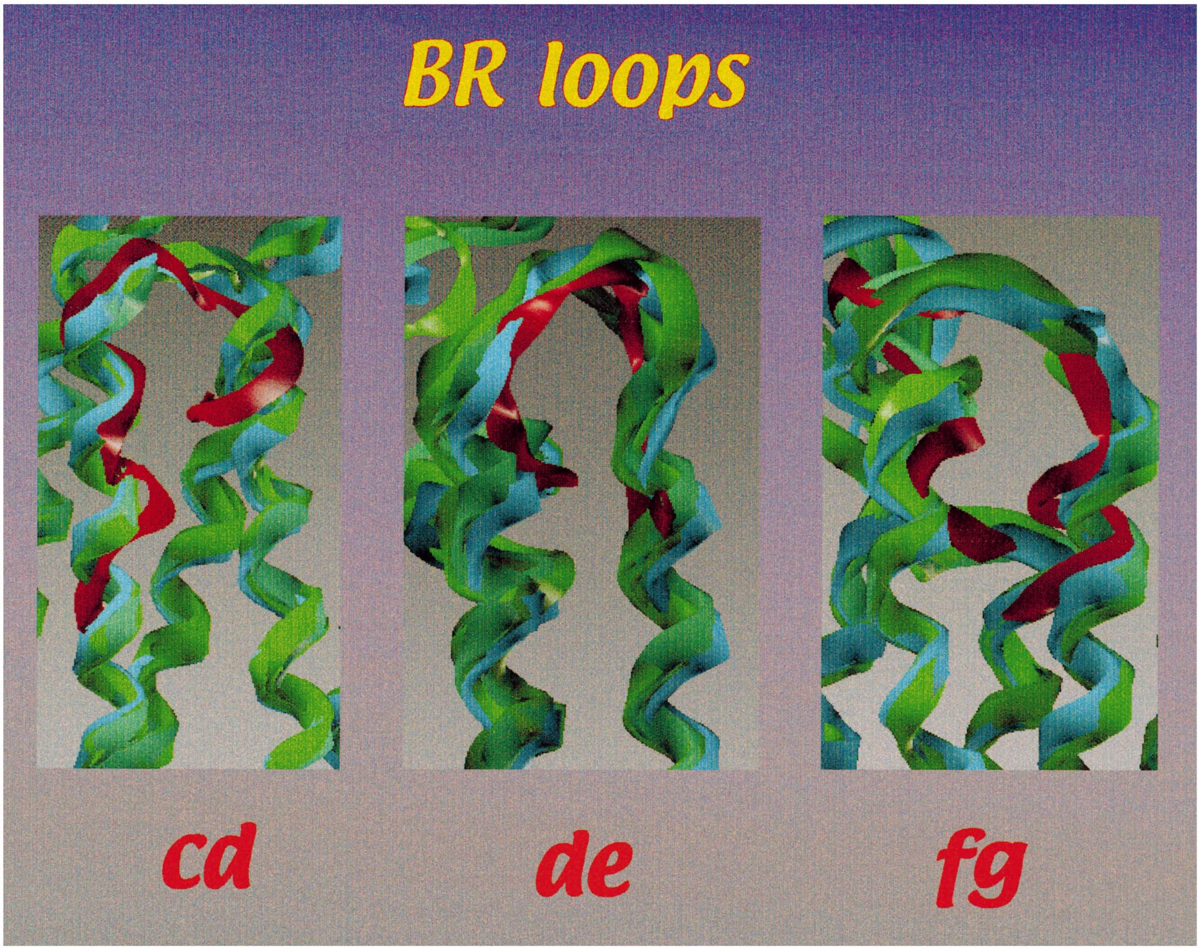


Fig. 4. Four crystal structures of bacteriorhodopsin were superimposed and represented in cyan and green [9–12]. The structures of the individual loops solved in this work are presented in red. Superposition of the loop structures on the crystal structures shows that the NMR structures fall within the envelope of the conformations from crystal structures of bacteriorhodopsin.

rmsd of 2.5. This includes the turn and the carboxyl-terminal helix of the peptide, which agree relatively well with the crystal structure according to the criteria described above. Superposition of the peptide on the crystal structure using residues 178–185 (bR) gave an rmsd of 2.0. Therefore, the amino-terminal helix of the peptide agrees well with the crystal structure. The fit is not as good if the whole range of ordered residues in the peptide structure are superimposed on the crystal structure. As in the case of DE, that is likely due to the paucity of available constraints between the helices. However, Fig. 3 shows that such a superposition actually gives relatively good agreement with the crystal structure and the solution structure of this peptide is also within the envelope of conformations represented in the available crystal structures, as shown in Fig. 4.

Residues 10–21 (99–110 of bR) of the CD loop peptide were superimposed on the crystal structure with an rmsd of two. The loop is centered on residues 96–102, so this superposition includes part of the loop and the carboxyl-terminal helix of the peptide. The amino-terminal helix also superimposes well: residues 91–96 superimpose with an rmsd of 2.1. As with the other loops, the fit is not as good if the whole range of ordered residues in the peptide structure are superimposed on the crystal structure, and as in the case of DE and FG, that is likely due to the paucity of available constraints between the helices. Fig. 4 shows that this structure of the CD loop is within the envelope of conformations represented in the available crystal structures.

A remarkable similarity in structure is observed between the individual peptides and the region of the parent protein with the same amino acid sequence. For each peptide, not only is a turn observed, but also the turn is of similar size to that observed in the crystal structure of the whole protein. Furthermore, in agreement with the structure of the whole protein, these subdomains appear to be helix-turn-helix motifs.

These results suggest that the loop regions connecting transmembrane helices in a protein like bacteriorhodopsin are stabilized largely by relatively short range interactions that can be found in a relatively small peptide as well as in the protein. Several recent studies on turns in G-protein-coupled receptors have shown that peptides containing sequences hypothesized to form turns connecting transmembrane helices show solution structures with a turn [6,7]. The results here for bacteriorhodopsin suggest that those turns from the G-protein-coupled receptors may be reporting accurately the secondary structure found in the intact receptor.

These observations raise some interesting issues on the factors leading to the stability of the bacteriorhodopsin structure. Helix–helix interactions have been shown to be very important to the stability of bacteriorhodopsin [8]. The results here suggest that at least some of the loop regions may also contribute to structural stability of the protein and/or play a role in the folding of the protein during/after synthesis. Since the peptides containing the sequences of these three loops of bacteriorhodopsin form turns independently of the remainder of the protein, the amino acid sequences for each of these regions must encode a loop structure. The short range interactions confined within these peptides are sufficient to stabilize the loop structure and thus may contribute to overall protein stability. The three loop regions studied here are well-defined in the crystal structures of bacteriorhodopsin. The conclusions for these three loops may not apply to the other loops of bacteriorhodopsin that are not well-defined in the crystal structures.

Acknowledgements

We thank Dr. A. Albert for helpful discussions and reading the manuscript and Dr. R. Wollman for collecting the NMR data. This work was supported by a Grant from the NIH (EY03328).

References

[1] J.F. Hunt, T.N. Earnest, O. Bousche, K. Kalghatgi, K. Reilly, C. Horvath, K.J. Rothschild, D.M. Engelman, Biochemistry 36 (1997) 15156–15176.

[2] A. Kumar, R.R. Ernst, K. Wuthrich, Biochem. Biophys. Res. Commun. 95 (1980) 1–6.

[3] K. Wüthrich, M. Billeter, W.J. Braun, J. Mol. Biol. 169 (1983) 949–961.

[4] H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, J.K. Lanyi, J. Mol. Biol. 291 (1999) 899.

[5] M. Dorey, P.A. Hargrave, J.H. McDowell, A. Arendt, T. Vogt, N. Bhawasar, A.D. Albert, P.L. Yeagle, Biochim. Biophys. Acta 1416 (1999) 217–224.

[6] P.L. Yeagle, J.L. Alderfer, A.D. Albert, Biochemistry 36 (1997) 3864–3869.

[7] D.F. Mierke, M. Royo, M. Pelligrini, H. Sun, M. Chorev, J. Am. Chem. Soc. 118 (1996) 8998–9004.

[8] T.W. Kahn, D.M. Engelman, Biochemstry 31 (1992) 6144–6151.

[9] N. Grigorieff, T.A. Ceska, K.H. Downing, J.M. Baldwin, R. Henderson, J. Mol. Biol. 259 (1996) 393–421.

[10] E. Pebay-Peyroula, G. Rummel, J.P. Rosenbusch, E.M. Landau, Science 277 (1997) 1676–1681.

[11] H. Luecke, H. Richter, J. Lanyi, Science 280 (1998) 1934–1937.

[12] E. Gouaux, Structure 6 (1998) 5–10.